Distributed Resource Allocation in OFDMA-Based Relay Networks

Vom Fachbereich 18
Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt
zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Dissertation

von
Dipl.-Ing. Ulrich Christian Müller
geboren am 12.03.1979 in Andernach

Referent:	Prof. Dr.-Ing. Anja Klein
Korreferent:	Prof. Dr.-Ing. habil. Volker Kühn
Tag der Einreichung:	19. Oktober 2009
Tag der mündlichen Prüfung:	12. Februar 2010

D 17
Darmstädter Dissertation
Darmstadt 2010

Danksagung

Ich konnte diese Arbeit anfertigen dank der Unterstützung, die ich von vielen Personen erfuhr. Den Personen, die mich am meisten unterstützten, danke ich an dieser Stelle.

Ich möchte mich besonders bei meiner Betreuerin Professor Anja Klein bedanken. Sie unterstützte mich von den ersten Ideen bis zur Vollendung dieser Arbeit. Ihre Vorschläge und konstruktive Kritik halfen mir sehr.

Ich danke Professor Volker Kühn für die Übernahme des Korreferats. Es war mir eine Freude, ihm und seinen Mitarbeitern meine Arbeit in Rostock zu präsentieren.

Allen ehemaligen Kollegen des Fachgebietes Kommunikationstechnik danke ich für die gute Atmosphäre, die ich während meiner fünfjährigen Tätigkeit im Fachgebiet erleben durfte. Den wissenschaftlichen Mitarbeitern danke ich für die zahlreich geführten Diskussionen, die es mir immer wieder erlaubten, Lösungen auch für komplizierteste Probleme zu finden. Den Mitarbeiterinnen des Sekretariats und den Systemadministratoren danke ich für ihr Engagement, das einen reibungslosen Ablauf der alltäglichen Arbeit garantierte. Auch den Studenten, deren Studien- oder Diplomarbeiten ich betreuen durfte, gilt mein Dank.

Ich bedanke mich bei meinen Projektpartnern, die während meiner Promotionszeit bei Nokia Siemens Networks tätig waren. Die Zusammenarbeit mit ihnen half mir bei der Entwicklung zahlreicher Ideen.

Zu besonderem Dank bin ich meiner Familie verpflichtet. Meine Eltern unterstützten mich stets so, dass ich meine volle Kraft in diese Arbeit investieren konnte. Meine Partnerin Nadine ließ mir besonders in schwierigen Momenten die mir nötige Unterstützung zukommen. Sie zeigt mir stets, was das Wichtigste in meinem Leben ist.

Darmstadt, März 2010

Ulrich Christian Müller

Kurzfassung

Konventionelle Mobilfunknetzwerke, die aus Basisstationen und Endgeräte der Teilnehmer bestehen, sind hinsichtlich ihrer Abdeckung und Kapazität beschränkt. Als vielversprechende Lösung für beide Probleme werden Relaisnetzwerke angesehen. In einem Relaisnetzwerk sind Relaisstationen installiert, um Informationen von einer Basisstation zu einem Endgerät weiterzuleiten. In einem Relaisnetzwerk wird typischer Weise eine Kombination von Zugriffsverfahren verwendet, wobei die Kombination aus Orthogonal Frequency Division Multiple Access (OFDMA), Space Division Multiple Access (SDMA) und Time Division Multiple Access (TDMA) von besonderer Wichtigkeit für zukünftige Netzwerke ist. OFDMA ermöglicht die Aufteilung des Spektrums in Zeit-Frequenz-Einheiten, die in Zeit- und Frequenzbereich definiert sind. In Kombination mit SDMA werden mehrere Strahlen bei einer Zeit-Frequenz-Einheit angewandt, um Zeit-Frequenz-Einheiten im Raum mehrfach zu nutzen. Eine Zeit-Frequenz-Einheit und ein Strahl bilden einen Ressourcenblock. Leistung wird einem Ressourcenblock zugeteilt, um ein Modulationsverfahren und eine Fehlercodierung anzuwenden. TDMA ermöglicht, dass Basisstation und Relaisstationen einer Zelle im Relaisnetzwerk ihre Übertragung koordinieren. Dazu werden Zeiteinheiten, die Slots genannt werden, der Basisstation und den Relaisstationen zugewiesen.

Das verfügbare Spektrum muss auch in einem Relaisnetzwerk effizient ausgenutzt werden, da Funkfrequenzen wertvoll sind. Die effiziente Nutzung des Spektrums ist besonders in Downlink-Richtung wichtig, da das Verkehrsaufkommen asymetrisch auf Uplink- und Downlink-Richtung verteilt ist. Die vorliegende Arbeit behandelt die Zuteilung der Ressourcen Strahlen, Ressourcenblöcke, Leistung und Slots in Downlink-Richtung eines Relaisnetzwerkes, um das Spektrum effizient zu nutzen. Ein Systemmodell wird eingeführt, um die Allokation der Ressourcen in einer Zelle eines Relaisnetzwerkes zu beschreiben. Das Systemmodell ist anwendbar auf zwei Arten von Szenarien. Die Arten von Szenarien unterscheiden sich im Zugriffsverfahren, das das Senden der Basisstation und der Relaisstationen in einer Zelle organisiert. In der erste Art von Szenarien sind Datenraten durch Rauschen limitiert, da ein orthogonaler Kanalzugriff vorausgesetzt wird. In der zweiten Art sind die Datenraten durch Gleichkanalinterferenz limitiert, da ein nicht-orthogonaler Kanalzugriff verwendet wird.

Zwei Ressourcen-Allokations-Probleme werden basierend auf dem Systemmodell definiert, wobei die Allokation in jeder Zelle einzeln betrachtet wird. Die Definitionen sind bezogen auf zwei Zielsetzungen. Die erste Zielsetzung ist die Maximierung der minimalen Datenrate, um eine faire Allokation im Sinne gleicher Datenraten pro

Endgerät zu erzielen. Die zweite Zielsetzung ist die Maximierung der Summe der Datenraten, wobei gleichzeitig jedem Endgerät eine minimale Datenrate zugesichert wird. Die Lösung der Ressourcen-Allokations-Probleme wirft für ein Relaisnetzwerk neue Fragen im Vergleich zu einem konventionellen Mobilfunknetzwerk auf. Es ist offen, welche Beiträge zur Lösung die Basisstation und die Relaisstationen in einer Zelle erbringen. Es ist offen, wie die Lösung zwischen Basisstation und Relaisstationen koordiniert wird, so dass die nötige Signalisierung gering bleibt. Es ist offen, wie eine Lösung mit geringem Rechenaufwand gefunden wird.

In der vorliegenden Arbeit wird motiviert, dass eine optimale Lösung der Probleme unter praktischen Gesichtspunkten nicht gefunden werden kann. Um dennoch machbare Lösungen zu finden, wird das Distributed Concept for Orthogonal Medium Access und das Distributed Concept for Reuse Medium Access eingeführt. Jedes Konzept wurde erstellt für eine Art von Szenarien, die im Systemmodell berücksichtigt werden. Jedes Konzept ist anwendbar auf beide Zielsetzungen. Jedes Konzept zerlegt ein Ressourcen-Allokations-Problem in kleinere Teilprobleme, so dass ein geringerer Rechenaufwand benötigt wird, um die Teilprobleme zu lösen. Die Teilprobleme werden zum Teil von der Basisstationen und zum Teil von den Relaisstationen gelöst, um den Rechenaufwand zu verteilen und um die Signalisierung gering zu halten.

Die Teilprobleme der beiden Konzepte werden als diskrete Optimierungsprobleme formuliert, da die Anzahl der Strahlen, der Ressourcenblöcke und der Slots durch eine natürliche Zahl gegeben ist. Selbst die Leistung, die eigentlich eine kontinuierliche Größe ist, kann nur eine diskrete Anzahl an Zuständen annehmen, da eine endliche Kombination aus Modulationsverfahren und Fehlercodierungen vorausgesetzt wird. Neue, adaptive Algorithmen, die eine adaptive Allokation bei geringem Rechenaufwand ermöglichen, werden eingeführt. Jedoch können diese adaptiven Algorithmen nur eingesetzt werden, wenn der Rechenaufwand von einer Basisstation oder Relaisstation erbracht werden kann. Falls dies nicht der Fall sein sollte, erlaubt jedes Konzept, dass ein adaptiver Algorithmus durch einen nicht-adaptiven Algorithmus ersetzt wird. Die nicht-adaptiven Algorithmen werden ebenfalls in dieser Arbeit vorgestellt.

Die Konzepte und die Algorithmen werden in einem beispielhaften Szenario untersucht. Es wird gezeigt, dass die Konzepte eine anwendbare und effiziente Allokation von Ressourcen in einem Relaisnetzwerk ermöglichen. Zusätzlich wird gezeigt, dass die adaptiven Algorithmen wesentlich bessere Ergebnisse erzielen als die nicht-adaptiven Algorithmen. Der dazu erforderliche Rechenaufwand kann von heutigen Prozessoren bereits erbracht werden.

Abstract

Conventional cellular networks consisting of Base Stations (BSs) and User Equipments (UEs) are limited in their coverage and capacity. Relay networks in which Relay Stations (RSs) are introduced to forward information from a BS to a UE are considered as a promising solution to both problems. In a relay network, a combination of multiple access schemes is typical, where the combination of Orthogonal Frequency Division Multiple Access (OFDMA), Space Division Multiple Access (SDMA) and Time Division Multiple Access (TDMA) is particularly important for future networks. OFDMA enables that the frequency spectrum is divided in time-frequency units defined in time and frequency domain. In combination with SDMA, multiple beams are applied to a single time-frequency unit in order to reuse time-frequency units in space. A time-frequency unit and a beam form a resource block. Power is allocated to a resource block in order to apply a modulation and coding scheme. TDMA ensures that a BS and the RSs of a cell in a relay network coordinate their transmissions. A number of time intervals called slots are allocated to the BS and to RSs for their transmissions.

The available frequency spectrum must be utilized efficiently even in a relay network since the frequency spectrum is expensive. The efficient usage of the frequency spectrum is particularly important in downlink direction due to an asymmetric traffic load distribution between uplink and downlink direction. This thesis deals with the allocation of resources namely beams, resource blocks, power and slots in the downlink direction in a relay network in order to utilize the frequency spectrum efficiently. A system model is introduced to describe the allocation of resources in a cell of a relay network. The system model is applicable to two types of scenarios differing in the medium access required to organize the transmissions of BS and RSs of a cell. In the first type of scenarios, the data rates are limited by noise since an orthogonal medium access is considered. In the second type of scenarios, the data rates are limited by co-channel interference since reuse medium access is considered.

Based on the system model, two resource allocation problems are defined, where the allocation is treated for each cell separately. The definitions are related to two objectives in order to provide a fair allocation in terms of data rates per UE and a high performance in terms of the sum of data rates. The first objective is the maximization of the minimum data rate and the second one is the maximization of the sum of the data rates subject to a minimum data rate provided to each UE. Concerning the solution of the resource allocation problems, new questions arise for a relay network

compared to a conventional cellular network. It is open which contributions of the solution are made by the BS and by the RSs of the cell. It is open how the solution is coordinated among BS and RSs such that the signalling overhead is kept low. It is open how a solution is found with limitations concerning the computational complexity.

It is motivated in this thesis that an optimum solution of the defined problems is impractical. In order to provide an applicable solution, the distributed concept for orthogonal medium access and the distributed concept for reuse medium access are introduced. Each concept is designed for one of the two medium access schemes considered in the system model. Each concept is applicable to both objectives. A concept decomposes a considered resource allocation problem in smaller subproblems such that lower computational complexity is required to solve the subproblems. The subproblems are partly solved by the BS and partly by the RSs in order to distribute the computational complexity and in order to keep the signalling overhead low.

The subproblems related to both concepts are formulated as integer programs since the numbers of beams, resource blocks and slots are integers. Even the power, which is actually a continuous quantity, can only reach an integer number of possible values since an integer number of modulation and coding schemes is assumed. Novel, adaptive algorithms enabling an adaptive allocation at a low computational complexity are presented for the subproblems. However, an adaptive algorithm is only applied, if the computational complexity can be fulfilled by a BS and RS. If this is not the case, each concept allows to replace an adaptive algorithm by the corresponding non-adaptive one also defined in this thesis.

The concepts including the introduced algorithms are evaluated in an exemplary scenario. It is shown that the concepts allow an applicable and efficient allocation of resources in a relay network. Additionally, the adaptive algorithms show strong performance gains compared to the non-adaptive ones. The computational complexity required to apply all adaptive algorithms can be offered by today's processors.

Contents

Chapter 1
Introduction

1.1 OFDMA-Based Relay Network

The radio access of today's cellular networks such as Global System for Mobile communications (GSM) [3GP00], Universal Mobile Telecommunications System (UMTS) [3GP09] or Institute of Electrical and Electronics Engineers (IEEE) 802.16 networks [IEE04] are based on the same topology. Typically, radio transmissions are applied between two kinds of stations: User Equipment (UE) and Base Station (BS). A BS enables the communication between UEs or provides the UEs a gateway to a wired backbone enabling an access of the UEs to other wired or wireless networks. A BS serves multiple UEs in an area called cell. These so called conventional networks are limited in two senses:

- A coverage problem exists. The average received signal power is a monotonically decreasing function of the distance between transmitting and receiving station [Par00]. Each data detection is disturbed by noise. A receiving station being too far away from the transmitting one cannot detect data since the signal power related to the noise power becomes too weak. This effect is intensified if the attenuation of the signal is increased further since the receiving station is shadowed by an obstacle like a building or a tunnel.
- A capacity problem exists. Transmissions take place in order to provide services. The data rate achieved for a single transmission is limited by the available bandwidth and power since noise and interference exist [Pro95]. Since multiple transmissions compete for the bandwidth and power in a conventional network, only a fraction of the bandwidth and power is available for a single transmission. If this fraction is too small, the data rate achieved by the transmission prevents that services are not provided in the conventional network.

Several solutions are conceivable but only partly applicable. The coverage problem can be solved if the power is increased by the transmitting station. This solution is not applicable since the maximum transmit power is restricted in order to protect co-existing networks from interference [ETS03] and a large radiated power is declined by customers due to concerns about health [Lin03]. The capacity problem may be solved if the bandwidth available for the conventional network is increased, but additional bandwidth is typically only available at higher carrier frequencies. For instance, today's GSM and UMTS networks operate with carrier frequencies of approximately

1 GHz or 2 GHz while the carrier frequencies of up to 5 GHz are proposed for next generation cellular networks [IST07a]. Since the attenuation of a signal is increased, the use of higher carrier frequencies leads to the coverage problem. The coverage and capacity problems can be solved if more BSs are deployed. More BSs provide a better coverage and larger data rates if the additional BSs reuse the bandwidth. However, each additional BS increases the expenditure related to the conventional network.

A solution related to both problems and attracting a strong interest in the last years is the introduction of a Relay Station (RS) as a third type of stations in a cellular network [WQDT01, PWS+04, CH04]. BSs, RSs and UEs form a relay network. A RS receives data actually not addressed to itself and forwards the data to the actually addressed station. The cell of a BS is enhanced by multiple RSs. A UE and a BS communicate directly or via multiple hops with the help of RSs. If multiple hops are used, a chain of transmitting and receiving stations is established. The source (e.g., a BS in downlink direction) transmits to an RS. Depending on the length of the chain, the data originally transmitted by the source is forwarded from RS to RS. The last RS forwards to the sink (e.g., a UE in downlink direction). The introduction of RSs improves the coverage of a BS [EVW01]. UEs located at the cell edge or shadowed by a building are able to communicate with the BS via a RS. The introduction of RSs improves the capacity since RSs additionally deployed in a cell increase the reuse of the available bandwidth [BYFP04, CH04, WQDT01, LLW+02]. For instance, a RS can provide radio access within a building without generating strong interference in neighbored cells. Beside a solution to the coverage and capacity problem, RSs have the advantage that a more cost-efficient deployment is possible for a relay network than for a conventional network as shown in [Tim05, MKWK07a, MSJF07]. Since a RS is not a gateway to the wired backbone and does not serve as many UEs as the BS, its functionality and size is reduced compared to a BS. Expenditures related to site rental, hardware, wired backbone and consumption of electricity are reduced. Hence, the cost related to a RS is smaller than the cost related to a BS.

Nevertheless, there are some drawbacks related to the usage of RSs. While conventional networks are operated successively for decades of years, the experience concerning the operation of relay networks is rather low. However, this drawback will vanish in the future. For instance, the specifications of relay networks in current standards such as Long Term Evolution (LTE) [3GP06] and IEEE 802.16j [IEE08] and the research related to relay networks will make their contributions to the required experience. A second drawback is that each hop introduces a delay. The data must be received by an RS before being forwarded. This drawback is weakened by an appropriate protocol defining the medium access in a relay network such that the data is forwarded immediately and by limiting the number of hops. The main drawback of the introduction of RSs is that each hop requires a transmission consuming bandwidth and power. This is not the case in a conventional network since multiple hops do not exist. Especially for relay networks, methods are required such that bandwidth and power are utilized efficiently.

This thesis aims at optimizing the usage of bandwidth and power in relay networks in order to make relay networks an even more applicable solution to the coverage and capacity problems. The thesis is related to a relay network defined by basic features as expected in next generation cellular networks such as LTE [3GP06] and IEEE 802.16j [IEE08], but the thesis is not related to a specific standard in order to cover a wide range of applications.

The relay network considered throughout this thesis is introduced in the following. Three types of stations exist: BSs, RSs, and UEs. The BSs and RSs are called Access Points (APs). The BSs are connected to a wired backbone. Each BS covers a cell in which several RSs and UEs exist. An exemplary relay network consisting of three cells is depicted in Fig. 1.1. As data services are essential in next generation cellular networks, an asymmetric distribution of the traffic load is expected. More data will be sent in downlink than in uplink direction because many data services are based on the client-and-server model [Tan96]. Since an optimized transmission is required more urgently in the downlink direction, this direction is considered in this thesis. All RSs are part of the infrastructure and deployed at a fixed position. Since the RSs are fixed, the link is optimized between BS and RS such that a Line-Of-Sight (LOS) link is given between these two kinds of stations. Hence, a low attenuation of signals transmitted from a BS to a RS is ensured. Additionally, the deployment of the RSs is optimized such that a sufficient coverage is achieved. Several strategies which can be pursued by the RS for the transmission exist. The most popular strategies are amplify-and-forward and decode-and-forward [LTW04,ZHF04,KGG05]. With amplify-and-forward, the received, noisy signal is amplified by the RS and forwarded. The main advantages of this strategy is that only a low computational complexity is required and that the RS does not introduce bit errors since the signal and the noise are amplified instead of detecting the signal. With decode-and-forward, the received signal is decoded, re-encoded and forwarded. This strategy allows the RS to change the modulation scheme and error correcting code in order to adapt the transmission to the wireless channel, but when the RS decodes the signal, bit errors may occur. Since the amplitude of the received and transmitted signals differ typically in several magnitudes, a simultaneous reception and transmission requires an extremely complex and expensive hardware. Hence, an RS operates always in a half-duplex mode [RW07], i.e, either it receives or it

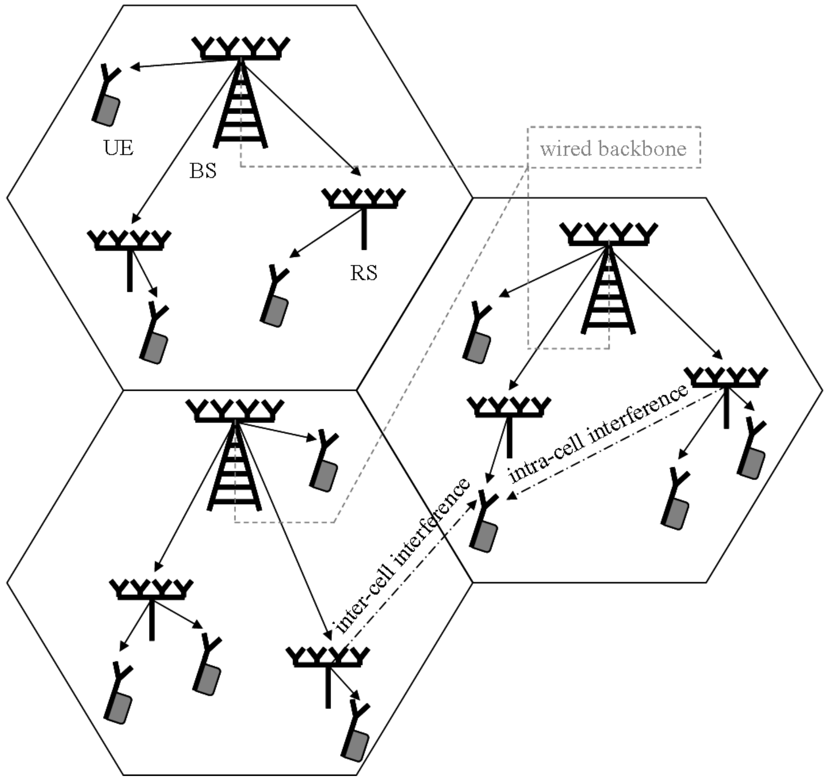

Figure 1.1. Exemplary relay network consisting of three cells.

transmits. The amplify-and-forward and decode-and-forward strategies are applicable to a non-cooperative and cooperative transmission:

- In a non-cooperative transmission, the information of the source (e.g., the BS transmitting in the downlink) is brought to the sink (e.g., the UE receiving in the downlink) over a chain, where the chain includes at least one RS [ZHF05]. The advantage of the non-cooperative transmission is that the path loss between source and sink is reduced compared to a direct transmission from source to sink [ZHF05].

- In a cooperative transmission, multiple transmitting stations coordinate their transmissions and the receiving station combines the signals received by the co-operating stations. A famous example is the relay channel consisting of three stations and analyzed in [vdM71, CG79]. The source and a RS cooperate and transmit to the sink. Amplify-and-forward and decode-and-forward strategy allow to exploit spatial diversity introduced by the usage of the RS [LTW04]. For the decode-and-forward strategy, various error correcting codes, encoding algorithms and decoding algorithms are proposed in [SE04, JHHN04, HN04]. A spatial diversity gain is also achieved if multiple RSs cooperate in a virtual antenna array [DSA02]. The RSs cooperate by using a distributed space time code [LW03, YSL06, JH06].

In this thesis, only non-cooperative transmissions are considered since the non-cooperative transmission can be integrated in a conventional network quite easily [HK06]. The number of hops is at most two since more hops cause a larger delay and allow an efficient usage of bandwidth and power only at a rather low Signal-to-Interference plus Noise Ratio (SINR) region [FY05]. The decode-and-forward strategy is applied such that the transmission can be adapted to the different qualities of the wireless channels related to the BS-to-RS link and to the RS-to-UE link. Additionally, the decode-and-forward strategy allows a demultiplexing of the data received by an RS if a single RS serves multiple UEs. In total, two types of connections exist in order to allow communication between a BS and a UE. The first one is the direct connection which consists of a BS-to-UE link. The second one is the two-hop connection consisting of a BS-to-RS and RS-to-UE link. A procedure is required for deciding whether a direct connection or a two-hop connection via one of the RSs is used. Since the number of hops is limited to two, no complex routing procedure is required. Various simple procedures differing in the applied metric are listed in [CFRZ07, MKWK07a]. For instance, UEs are assigned to the BS or an RS according to the shortest distance [CFRZ07] or according to the highest expected data rate [CFRZ07, MKWK07a].

The APs of a cell provide the data to the UEs. A frequency spectrum limited in bandwidth and identified by the carrier frequency is assigned to the APs of the cell. If the carrier frequency is reused by several APs, transmissions are disturbed by co-channel interference. If a reused is applied among the APs of the same cell, the co-channel interference is called intra-cell interference. If a reused is applied among the APs of different cells, the co-channel interference is called inter-cell interference. Since an AP may serve multiple receiving stations and multiple APs transmit in a cell, a multiple access scheme is required. In this thesis, a combination of Orthogonal Frequency Division Multiple Access (OFDMA), Time Division Multiple Access (TDMA) and Space Division Multiple Access (SDMA) is considered.

OFDMA is a multiple access scheme applied in various wireless systems like IEEE 802.11 [IEE07] or IEEE 802.16 [IEE04] and is a promising candidate for next generation cellular networks, e.g., accepted for LTE [3GP06] and IEEE 802.16j [IEE08]. OFDMA employs Orthogonal Frequency Division Multiplexing (OFDM) [NP00, LL05]. OFDM divides the total available bandwidth in small bands called subcarriers. Each subcarrier is loaded with data symbols, where a guard interval is introduced between two consecutive data symbols. The subcarriers are orthogonal to each other and do not interfere with each other in a multipath propagation scenario if the guard interval is designed well [NP00, LL05]. A first advantage of OFDM is that subcarrier can be allocated different power values such that different number of bits are allocated to subcarriers as known from the water filling algorithm [PNG03, PF05]. However, power values are not mapped to bits by a continuous function since a finite set of modulation and coding schemes exists, where a modulation and coding scheme is defined as the combination of a scalar linear modulation scheme [Pro95] like Phase Shift Keying (PSK) or Quadrature Amplitude Modulation (QAM) and an error correcting code, e.g., a convolutional code [Bos98]. A second advantage of OFDM is that disjoint sets of subcarriers can be formed. OFDMA is based on this partitioning [NP00, LL05]. An advantage of OFDMA is the flexibility concerning the allocation of the resources power and subcarriers in a multiuser environment [NP00, LL05]. Based on channel quality information (CQI) describing the instantaneous quality of the wireless channel, an adaptive allocation of power and bits as enabled for OFDM is also applicable in a multiuser environment. Additionally, the set of subcarriers allocated to a receiving station can be designed adaptively.

Since multiple APs transmit in a cell in downlink direction, the question arises how the APs access to the wireless channel. APs deployed closely interfere with each other if they use the same subcarriers simultaneously. In this thesis, TDMA is applied in order to avoid this interference. APs transmit in a frame. A frame consists of several time units called slots, where a slot is the smallest time unit allocated to an AP. Each AP uses a set of consecutive slots for its transmission. TDMA is chosen because of the following reasons:

- A time multiplexing is required anyway since a RS operates in half-duplex mode.
- TDMA enables that an AP uses all subcarriers being available in a cell. An allocation of subcarriers taking into account CQI is expected to achieve a higher data rate per receiving station if the number of available subcarriers is large due to the frequency selective nature of the wireless channel [LL05].
- TDMA enables an easy signaling when an AP starts and ends its transmission. Since an AP transmits in consecutive slots, only the index values of first and last slots must be signalled to an AP.
- TDMA is of practical interest. For instance, it is part of the IEEE 802.16j [IEE08] standard.

The combination of OFDMA and TDMA allows the definition of time-frequency units. A time-frequency unit is defined in time domain by a set of successive slots and in frequency domain by a set of adjacent subcarriers.

SDMA allows the reuse of time-frequency units in space [PNG03, ST81]. If the transmitting AP is equipped with an antenna array, its elements are controlled in amplitude and phase such that a beam is formed. The information about amplitude and phase is represented by a beamforming vector. Multiple beamforming vectors applied to the same time-frequency unit allow that an AP transmits in different directions and enable a spatial multiplexing. If the channel state information (CSI) defined as the instantaneous attenuation and phase shift caused by the wireless channel is known to the transmitting AP, various beamforming techniques leading to high data rates exist [PNG03, Qiu04, SSH04, JUN05]. However, the CSI is hard to obtain for the APs [VTL02]. Alternatives which enable SDMA although beamforming vectors are created without instantaneous information about the channel are proposed in literature [VTL02, VALP06, GRB06a, RGT08]. These alternatives are presented more deeply in Section 1.3. One of these alternatives is considered in this thesis. The APs choose beamforming vectors out of a set of pre-defined vectors which are created before the operation of the relay network and without instantaneous information about the channel. The chosen vectors form a set, where such a set is formed for each time-frequency unit used by an AP. The beams chosen for a time-frequency unit used by an AP is called grid of beams throughout this thesis. A time-frequency unit and a chosen beam form a resource block. Resource blocks are allocated to a link between the transmitting AP and the receiving station. An adaptive allocation of bits is possible for the resource blocks. This means that different modulation and coding schemes are applied for different resource blocks. The detection of data carried by a resource block is disturbed by noise, co-channel interference and interference which occurs between resource blocks transmitted by the same AP and using the same time-frequency unit. This latter interference is called inter-beam interference.

1.2 Resource Allocation Problems

If information about the wireless channel, like CQI or CSI, is available at the transmitting APs, various resource allocation problems related to the combination of OFDMA, TDMA and SDMA occur in the relay network in which all direct and two-hop connections are established. The objectives, the constraints linked with these objectives, the resource allocation problems themselves and the interdependencies of these problems are presented in non-mathematic terms in this section. For the sake of simplicity, a single cell of the relay network is considered. This assumption is generalized in the end of this section.

The definition of the objectives are related to a trade-off. Like in every cellular network, the trade-off exists between fairness and performance in the relay network [SAR09]. In order to reveal this trade-off, the term user rate is defined. The user rate is defined as the number of bits received by a UE from the BS in a frame. The bits are transmitted via a direct or a two-hop connection. Since bandwidth and power are limited, it is impossible to maximize the user rate of all UEs. The solution of a resource allocation problems leads either to a fair or to a strong performing result. Fairness means in this context that all UEs shall achieve the same user rate. Performance is represented by the sum rate defined as the sum of all user rates. Corresponding to the extreme cases of the trade-off, two objectives are considered in this thesis:

- The fair objective is the maximization of the user rate which corresponds to the UE which has the minimum user rate among all UEs. This objective is fair since each UE achieves the same user rate in the ideal case that the number of bits allocated to a resource block is a continuous function of the achieved SINR. However, this objective leads to a weak performance. If a single UE suffers from bad channel conditions, this UE will require a large amount of bandwidth and power and the remaining UEs achieve only a low user rate due to the lack of resources.
- The objective leading to a strong performance is given by the maximization of the sum rate, but this leads to an unfair solution since some UEs achieve large user rates while the remaining ones achieve only low user rates.

The objectives are coupled with two constraints in order to satisfy the Quality of Service (QoS) demanded by the UEs. The first constraint claims that a transmission is successful. It must be ensured for both objectives that the transmitted bits are provided with a satisfying bit error probability value. The second constraint is introduced in order to ensure a minimum of fairness. Since the fair objective leads to a fair result, this constraint is only applied in combination with the objective leading to a strong performance. The constraint is that each UE is guaranteed a minimum user rate. By varying the minimum user rate, each degree of fairness is achieved as long as the minimum user rate can be provided.

The objectives and the constraints linked with the objectives are considered if problems related to the allocation of resources are defined. Since the combination of SDMA, OFDMA and TDMA is treated in this thesis, three resource allocation problems exist, where each problem is related to a multiple access scheme. The resource allocation problems considered in this thesis and their interdependency are illustrated in Fig. 1.2. These problems are:

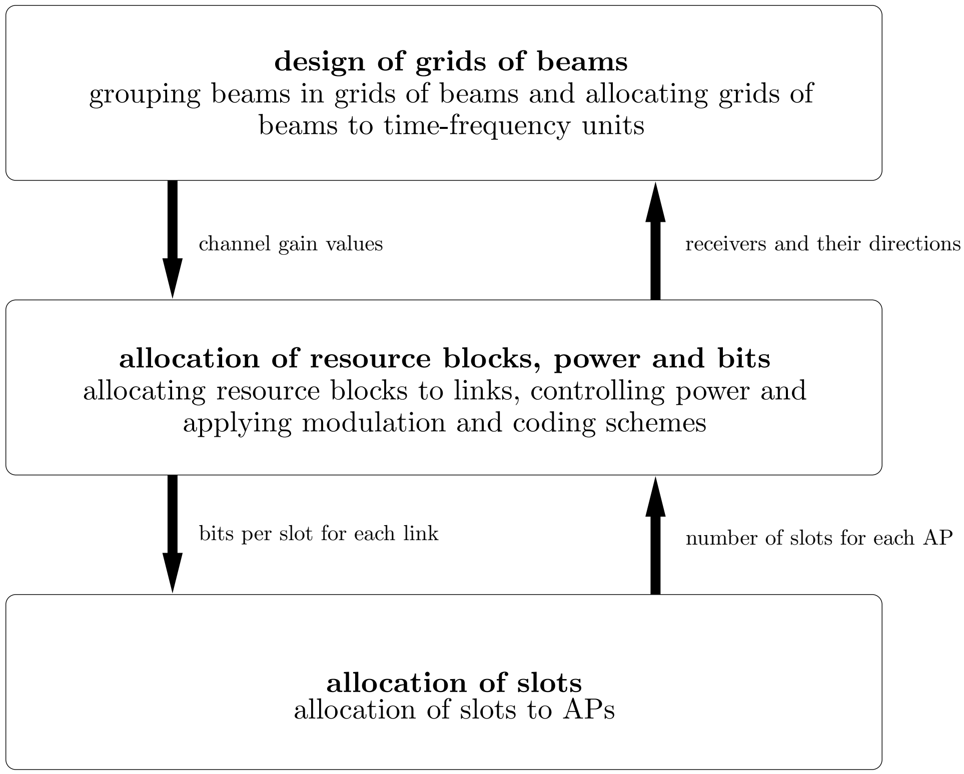

Figure 1.2. Resource allocation problems in the relay network.

- The problem related to SDMA is called design of grids of beams. It corresponds to the question which beams, where a beam is identified by its beamforming vector, are grouped in the same grid of beams and which grids of beams are allocated to time-frequency units. The solution must take into account that beams cause inter-beam and co-channel interference.

- The problem related to OFDMA is called allocation of resource blocks, power and bits. It corresponds to the questions which resource block is allocated to a link, which power is allocated to a resource block and which modulation and coding scheme is applied for a resource block. The solution must solve the conflict between stations served by the same AP and competing for resource blocks and power. Additionally, the solution must take into account that only a finite set of modulation and coding schemes is available in a relay network.

- The problem related to TDMA is called allocation of slots and corresponds to the problem which slots are reserved for an AP for its transmission. The solution must solve the conflict between transmitting APs competing for slots.

The resource allocation problems affect each other. The design of grids of beams requires two types of information. The first one implies which resource blocks are allocated to a receiving station. The second one implies the direction of the receiving station such that a beam can be steered in this direction. The latter information is included in the received power when multiple beamforming vectors are tested for the same receiving station and resource block. However, both types of information is only available if the problem of the allocation of resource blocks, power and bits is solved. The allocation of resource blocks, power and bits requires the channel gain values describing the wireless channel including antenna gains. The channel gain values result from the solution of the design of grids of beams. The allocation of resource blocks, power and bits is optimized for a two-hop connection if the bits transmitted on the BS-and-RS and RS-to-UE link are balanced such that a RS receives as many bits as it forwards. Hence, the allocation of resource blocks, power and bits is only solved if the number of slots allocated to the APs are known, but this is the result of the allocation of slots. However, the allocation of slots is based on the bits transmitted per slot for each link in order to balance the transmission of the APs such that a RS receives as much as it is able to forward. The conclusion is drawn that the three problems cannot be solved independently.

So far, the resource allocation problems are defined for a single cell. Although this restriction is initially made for the sake of simplicity, the restriction is of practical interests and motivated as follows. The solution of the resource allocation problems require information about the wireless channel like CQI or CSI. This information is time varying. Whenever the information changes, the resource allocation problems must eb solved. The solutions of the resource allocation problems defined among several cells lead to a large amount of signalling and a large computational complexity [KF08]. If each cell solves its resource allocation problem individually, signalling overhead and computational complexity is manageable. Hence, problems related to the radio transmissions are defined for each cell independently.

1.3 State-of-the-Art

In this section, an overview about previous work is given. At first, work related to the analysis of the relay network is presented. Then, work related to resource allocation problems dealing with OFDMA and SDMA in the conventional and the relay network is shown in order to integrate the resource allocation problems presented in Section 1.2 in previous work.

The relay network has various advantages compared to the conventional one [PWS+04]. As shown in the following, these advantages are related to solutions of the coverage and capacity problems. Works related to the coverage problem are presented in [EVW01, MKW06, MKW07, BYFP04]. In [EVW01], the coverage of a HIPERLAN system [ETS97] is improved by the usage of RSs. The coverage extension is analyzed for an IEEE 802.16 network [IEE04] in [MKW06, MKW07]. In [EVW01, MKW06, MKW07], the coverage of a BS is extended by UEs which act as RSs since they are willed to forward information of other UEs. In [BYFP04], the sum rate of a cell is determined in the relay network if the RSs are part of the infrastructure. Works related to the capacity problem are presented in [CH04, WQDT01, LLW+02], where a RS is always part of the infrastructure. In [CH04], it is shown that the sum rate is enhanced by the usage of RSs and simultaneously user rates are more equally distributed among the UEs compared to conventional networks. In [WQDT01, LLW+02], the strategy is pursued that the load of multiple cells is balanced among cells in the relay network. It is shown that improvements related to the sum rate, the coverage and the required transmit power make the relay network superior to a conventional network without any load balancing. The improvement of sum rate is only achieved if the number of hops is kept low [FY05]. Depending on the SINR, the optimum number of hops is given in [FY05]. The improved coverage and sum rate make the relay network more cost-efficient than the conventional network if the cost related to a RS is small enough. Based on models describing the cost of the conventional and relay network, the breakeven point where the relay network is more cost-efficient is determined in various scenarios in [Tim05, MKWK07a, MSJF07]. In each scenario, the relay network is only more cost-efficient if the available bandwidth is exploited efficiently.

Related to the considered relay network, three opportunities exist how the usage of the bandwidth is optimized. The first opportunity is related to the questions how bandwidth is reused among APs in the relay network and at which slots the RSs transmit and receive. The second opportunity is the optimization of the OFDM transmissions and the last one is the optimized combination of SDMA and OFDMA. Work related to these opportunities are shown in the following.

The questions how bandwidth is reused in the relay network and at which slots the RSs transmit and receive are addressed in [DXWS07, SAY06, SWPI03, UMK06]. In [DXWS07], the bandwidth is exploited efficiently by a proposed soft reuse. The soft reuse enables a frequency reuse of one among cells by applying various power masks varying in the time or frequency domain for the APs of the relay network. Since an RS operates typically in half-duplex mode, a single two-hop connection requires always two slots while a direct connection requires only one. Methods related to the reuse of the bandwidth within a cell are proposed in [SAY06, SWPI03, UMK06] in order to compensate for this loss. In [SAY06] multiple RSs are divided in two groups. While the RSs of the first group are receiving from the BS, the RSs of the second group forward to the UEs. In the next slot, the roles of the groups are changed. The advantage of this strategy is that the BS always transmits, but this is paid by co-channel interference. In [SWPI03], RS are grouped taking into account the co-channel interference among cells. Each group is allocated a slot of fixed duration. In [UMK06], the slots have a variable duration which is proportional to the number of UEs assigned to the group, but the optimized number of slots concerning any objective is not found so far. Hence, a solution of the problem of the allocation of slots is missing.

If OFDMA is used in the relay network, the bandwidth is exploited more efficiently if power and subcarriers are allocated adaptively if information about the wireless channel is available for the transmitting APs. Works related to an adaptive allocation exist in literature, both for the conventional and the relay network. In order to give a structured overview, the works related to the conventional network are reviewed at first, followed by the works related to the relay network.

The works related to the conventional network are divided in works related to single UE and multiple UEs. If a single UE receives from the BS in a cell, adaptive methods related to OFDM are applicable. The methods adapt the allocation of power and bits according to the frequency selective channel. The allocation of power and bits according to the water filling algorithm [PNG03, PF05] leads to the maximized number of bits transmitted per channel usage with an arbitrary small bit error probability. However, the water filling algorithm is based on a continuous function representing the transmitted number of bits depending on the allocated power, but in practice, only a finite set of modulation and coding schemes is available. With respect to a finite set, various power loading algorithms exist, e.g., proposed in [HH87, CCB95]. In [HH87], the power which is limited in sum is loaded to subcarriers such that the number of transmitted bits per channel usage are maximized. In [CCB95], power is loaded such that the allocated power is minimized subject to an achieved user rate.

If multiple UEs receive from the BS in a cell, an adaptive allocation of subcarriers to the UEs is possible. If the UEs suffer from mutually independent fading, large gains in terms of bit error probability, user rates or sum rate are possible due to an effect called multi-user diversity [LL05]. Multi-user diversity occurs when the UE providing the best channel condition is selected for the allocation out of multiple UEs. Subcarriers are allocated by a simple greedy algorithm such that the sum rate is maximized [LL05]. However, the user rates are not fairly distributed among the UEs in this case. A fair distribution is achieved if the objective is the maximization of the minimum user rate [ECV03]. Various methods exist where the adaptive allocation is combined with the consideration of different priorities of the UEs. These methods are known as opportunistic scheduling [LCS01, VTL02, GRB06a]. The jointly optimized allocation of subcarriers, power and bits is addressed in [WCLM99, ZL04, RC00, KPL06]. Since the number of subcarriers is discrete and the set of bits which can be allocated is finite, the problems describing a joint allocation are combinatorial ones and are hard to solve [NW99, LZ06]. In [WCLM99], the objective is the minimization of the allocated power while a user rate is being guaranteed for each UE. The problem is relaxed as a convex optimization problem and solved with large computational effort. In [ZL04], the allocation aims at maximizing the sum rate while a user rate is being guaranteed for each UE. If the guaranteed user rate is not achievable, UEs are assigned to neighbored cells. A fair objective in sense of equal user rates is treated in [RC00], where the allocation aims at the maximization of the minimum user rate. In [KPL06], the allocation of subcarriers, power and bits is formulated as linear integer program, i.e., the objective function and the constraints of the optimization problem are given by linear equations or inequations and the optimization variable is an integer variable. The formulation helps to find algorithms of low complexity, but provides only a suboptimum allocation of the integer programs.

For the relay network, only a few references exist dealing with adaptive allocation related to OFDM and OFDMA. The question how bandwidth and power are allocated among several hops of a single non-cooperative transmission is answered in [Doh04, SZQY05], where the objective is the maximization of the user rate. Allocation problems related to multiple UEs are addressed in [LHYT06, HRW+07, MKWK07b, LL06, MKR09]. In [LHYT06], the problem of allocating resource units which represent subcarriers or slots and the problem of finding the route between source and sink are formulated as a joint optimization problem for a scenario with multiple UEs and RSs. Since the problem has been proven to be Nondeterministic Polynomial time hard (NP-hard), an algorithm finding a suboptimum solution with a low computational complexity is presented. In [HRW+07], an algorithm is proposed which aims at maximizing the sum rate in a cell by allocating subcarriers and loading power adaptively in a cell of the relay network. In [MKWK07b], a resource allocation algorithm for the relay network is presented which minimizes the transmit power of the BS and the RSs subject to the constraint that each UE achieves a minimum data rate. The algorithms presented in [LHYT06, HRW+07, MKWK07b] require instantaneous information about all wireless channels at a central point in the cell, e.g., the BS. Since the information about all RS-to-UE links must be reported to the BS, a high signaling overhead is expected for these algorithms. The signaling overhead is avoided if the information about the RS-to-UE links is directly used by the RSs. In [LL06], an algorithm allocating subcarriers in a distributed manner is presented. Based on average channel state knowledge, a central unit (e.g., a BS) allocates the number of subcarriers to the RS. Each RS performs the power allocation to its subcarriers. In opposite to the constraints assumed in this thesis, a station can simultaneously receive and transmit on different subcarriers separated in frequency domain in [LL06]. The allocation presented in [LL06] is proportionally fair. Work aiming at maximizing the sum rate or achieving totally fair distribution of the user rates among the UEs is missing if the signaling overhead is kept acceptable. A first approach is addressed in our work [MKR09], where the allocation aims at maximizing the sum rate.

If the transmitting APs are equipped with multiple antennas, SDMA is possible [PNG03, ST81]. Especially, the combination of SDMA and an adaptive allocation of time-frequency units promises high user rates and a high sum rate in the downlink of OFDMA-based networks. Various concepts exist for this combination. At first concepts related to conventional networks, then related to the relay network are reviewed.

For conventional networks, the concepts differ in the information which must be known at the BSs about the channel states. If the BSs have CSI given by the instantaneous amplitudes and phases of the channels, transmit beamforming can be performed by a BS and several data streams intended to different receiving stations are sent on a single time-frequency unit [PNG03, Qiu04, SSH04, JUN05]. In [SS04, FBB+08, Mac09], joint concepts combining transmit beamforming and adaptive allocation of time-frequency units promise a large sum rate in a single cell. These concepts suffer from inter-cell interference in cellular networks. In [PIPNL04], transmit beamforming coordinated among BSs is proposed. The BSs cooperate in order to mitigate inter-cell interference. CSI about the channels between all receivers and all cooperating BSs is required. However, CSI between a single BS and multiple UEs is hard to obtain in a practical network, but obtaining the information between multiple BSs and multiple UEs is nearly impossible since an unacceptable amount of signalling is required. Alternatives exist which do not apply transmit beamforming requiring CSI. In [VTL02], an opportunistic beamforming and an adaptive allocation of time-frequency units is proposed. Beams are formed randomly on time-frequency units. A user determines the SINR value for each beam on a time-frequency unit and feeds these values back to the BS. The feedback is exploited to allocate the time-frequency units to proper users by a smart scheduling. In [VALP06], a concept of opportunistic beamforming and adaptive allocation of time-frequency units is proposed for multiple cells as an extension of [VTL02]. BSs cooperate in beamforming and allocation of subcarriers in order to reduce the inter-cell interference. However, the cooperation is rather complex since the concept is based on an optimization problem which is defined among several BSs and which must be solved whenever the SINR values change. In [GRB06a], a concept is proposed in which pre-defined beams realized by Chebyshev beam weights [ST81] are used. The beams are grouped in grids of beams such that each beam belongs to exactly one grid of beams. The beams are grouped such that inter-beam interference is minimized. After the decision which beam is applied on a time-frequency unit, pilots are transmitted by the BSs of all cells in the network simultaneously. Each user estimates the SINR for each beam on a time-frequency unit. The SINR values are fed back by the users to the BSs. The BSs perform an adaptive allocation of time-frequency units based on these SINR values. In the actual transmission, the same power and the same beams are applied on a time-frequency unit as for the pilot transmission in order to exploit precise SINR values in the smart scheduling algorithm. In [RGT08], the proposal is made that the use of beams is adapted to the distribution of the users in the cell leading to performance gains, e.g., in a hot-spot scenario. Since only one beam is considered in the grids of beams, it remains unclear how to group beams in grids of beams if a grid of beams consists of more than one beam. The main advantage of the concept presented in [GRB06a, RGT08] is that each cell applies adaptive allocation of time-frequency units independently from the other cells taking into account inter-beam and inter-cell interference. This makes the concept rather interesting for practical application.

Due to the results in conventional networks, the combination of SDMA and adaptive allocation of time-frequency units promises also good results in the relay network. Transmit beamforming, opportunistic beamforming and grids of beams are applicable for the relay network. Two additional aspects must be considered compared to the conventional network. At first, intra-cell interference exists in the downlink if two APs transmit at the same carrier frequency simultaneously. Additionally, the data rate of the BS-to-RS and the RS-to-UE link must be balanced. In [CZT07], UEs are grouped in spatially compatible groups which share the same time-frequency units. The spatial compatibility is determined based on the directions of arrival for the signals received by the UEs. Then, time-frequency units are allocated to the groups such that the sum rate is maximized or the user rates are distributed fairly. Both allocations aim at balancing the data rate on the BS-to-RS and RS-to-UE link. In [CFRZ08], a similar concept is proposed. UEs are grouped based on an average interference. The adaptive allocation of time-frequency units to single UEs is excluded in both works. Hence, a joint concept for the combination of SDMA and adaptive allocation of time-frequency units is missing for the relay network so far.

1.4 Problem Statement

Based on the description of an OFDMA-based relay network given in Section 1.1 and a comparison of the resource allocation problems defined in Section 1.2 and the state-of-the-art reviewed in Section 1.3, the problems addressed in this thesis are stated in this section.

Since each cellular network is affected by the coverage or capacity problem, two scenarios are of practical interest for the relay network:

- Scenario 1: in which user rate and sum rate values are limited by noise.
- Scenario 2: in which user rate and sum rate values are limited by co-channel interference.

Since the trade-off between fairness and performance exist in each cellular network, the resource allocation problems illustrated in Fig. 1.2 occur for all solutions of this trade-off. In order to consider this trade-off, the following objectives are treated in this thesis:

- Objective 1: maximize the minimum user rate.
- Objective 2: maximize the sum rate while each UE is guaranteed a minimum user rate.

If the minimum user rate is varied in the second objective, a wide range of solutions is covered for the trade-off between fairness and performance.

When the resource allocation problems illustrated in Fig. 1.2 and the review of the existing literature provided in Section 1.3 are compared, three problems are left open for each cell in the relay network. These problems are:

- Problem 1: Typically, more data are transmitted via a BS-to-RS link than a BS-to-UE or RS-to-UE link since a RS can serve multiple UEs. Additionally, UEs are not always distributed in a cell uniformly. Time-frequency units are exploited efficiently if beams are steered in the direction of the receiving UEs and more often in the direction of the RSs. If grids of beams are applied, it is open how often beams are used and how they are grouped in grids of beams in a cell of the relay network.

- Problem 2: Related to the allocation of resource blocks, power and bits, it is open how an adaptive allocation is performed among the APs of a cell without a large signalling overhead caused by bringing the required information to a central unit.
- Problem 3: It is open for how many time units represented by slots in this thesis the BS and RSs shall transmit, if multiple UEs are served and resource blocks, power and bits are allocated adaptively.

Although the description of the resource allocation problems is given in non-mathematical terms in Section 1.2 and some of the problems are formulated partly for the conventional network or partly for the relay network, precise and complete formulations are missing for the relay network. Formulations related to the scenarios and objectives introduced in this section ensure that the problems are of practical interests. The solutions of all three problems must consider the interdependencies of the problems. In order to make the solutions applicable, the computational complexity and the signaling overhead between the APs must be limited.

1.5 Contribution and Outline of the Thesis

This section reveals the main contributions and gives an overview about the structure of this thesis. The contributions and the structure are related to the problem statement introduced in Section 1.4.

In Chapter 2, a novel system model is introduced for a single cell of the relay network, where a cell consists of the BS, multiple RSs and multiple UEs. The system model describes the application of the multiple access schemes OFDMA, TDMA and SDMA and the application of a limited number of modulation and coding schemes. The system model is defined with respect to the scenarios in which user rate and sum rate values are limited by noise and by co-channel interference.

In Chapter 3, two joint formulations are given for the three resource allocation problems illustrated in Fig. 1.2. The first formulation is related to the objective of the maximization of the minimum user rate and the second one to the objective of the maximization of the sum rate. Two novel concepts are proposed in order to find a solution. Each concept defines a framework how the joint problems are decomposed in subproblems in order to make a solution applicable. The first concept is related to the scenario in which user rate and sum rate values are limited by noise, the second one to the scenario in which user rate and sum rate values are limited by co-channel interference. Both concepts are applicable to both objectives.

In Chapter 4, the subproblems resulting from the decomposition in the first concept are treated. The subproblems are related to the maximization of the minimum user rate and the maximization of the sum rate. Novel adaptive algorithms taking into account the available information about the position of the RSs and UEs and about the wireless channels are introduced in order to solve the subproblems. Non-adaptive algorithms are introduced and serve as trivial solutions of the subproblems.

In Chapter 5, the subproblems resulting from the decomposition in the second concept are treated, where the scenario in which user rate and sum rate values are limited by co-channel interference is considered. Again, the subproblems are related to the maximization of the minimum user rate and the maximization of the sum rate. Novel adaptive and non-adaptive algorithms are introduced.

In Chapter 6, the performance of both concepts is evaluated. The evaluation is related to both scenarios and both objectives considered in this thesis. Additionally, the signalling overhead and the computational complexity required for the application of the concepts is shown.

In Chapter 7, the main results of this thesis are summarized and the thesis is concluded.

Chapter 2
System Model

2.1 Introduction

This chapter gives the system model for a cell of the relay network. The chapter is organized as follows. Section 2.2 presents the topology of the cell. A mathematical representation of the links between stations is introduced. In Section 2.3, a model describing the medium access as the combination of OFDMA, TDMA and SDMA is presented. The model considers that several APs transmit in downlink direction and multiple UEs are served in the cell. The model covers two cases. In the first case, time-frequency units are used by an AP exclusively and the user rates and sum rate are limited by noise. In the second case, time-frequency units are reused among APs and co-channel interference exists. Furthermore, various resource units are defined, namely the resource units frame, slot, time-frequency unit and resource block. Section 2.4 gives the models of the transmit unit and the receive unit. These models are used to show the restrictions to which the resource allocation problems are subject. The models consider that a finite set of modulation and coding schemes exists. In Section 2.5, the signals received by a UE and an RS are described and the SINR values are derived. In Section 2.6, the performance measures representing the performance achieved by a single UE and by the cell as a whole are defined.

2.2 Topology of the Cell

This section presents the topology of the cell. The considered scenario is already introduced in Section 1.1 and depicted in Fig. 1.1. The cell consists of the BS, N_{RS} fixed RSs and N_{UE} UEs. The UEs are distributed randomly in the cell according to the Probability Density Function (PDF) $\rho(x, y)$, where x and y are Cartesian coordinates describing the area of the cell. A UE is connected to the BS by the mean of a direct or a two-hop connection. The topology defining which stations are linked in a cell of the relay network is given by a graph and illustrated in Fig. 2.1. Each station is given by a node of the graph. The graph has the structure of a tree with a depth of two edges. The UEs receive data from the BS by two types of connections. The first type is a direct connection since a BS transmits to a UE via a BS-to-UE link without an intermediate RS. The second type is a two-hop connection in which the BS first transmits to an RS via a BS-to-RS link and then the RS forwards the data to the UE via an RS-to-UE link.

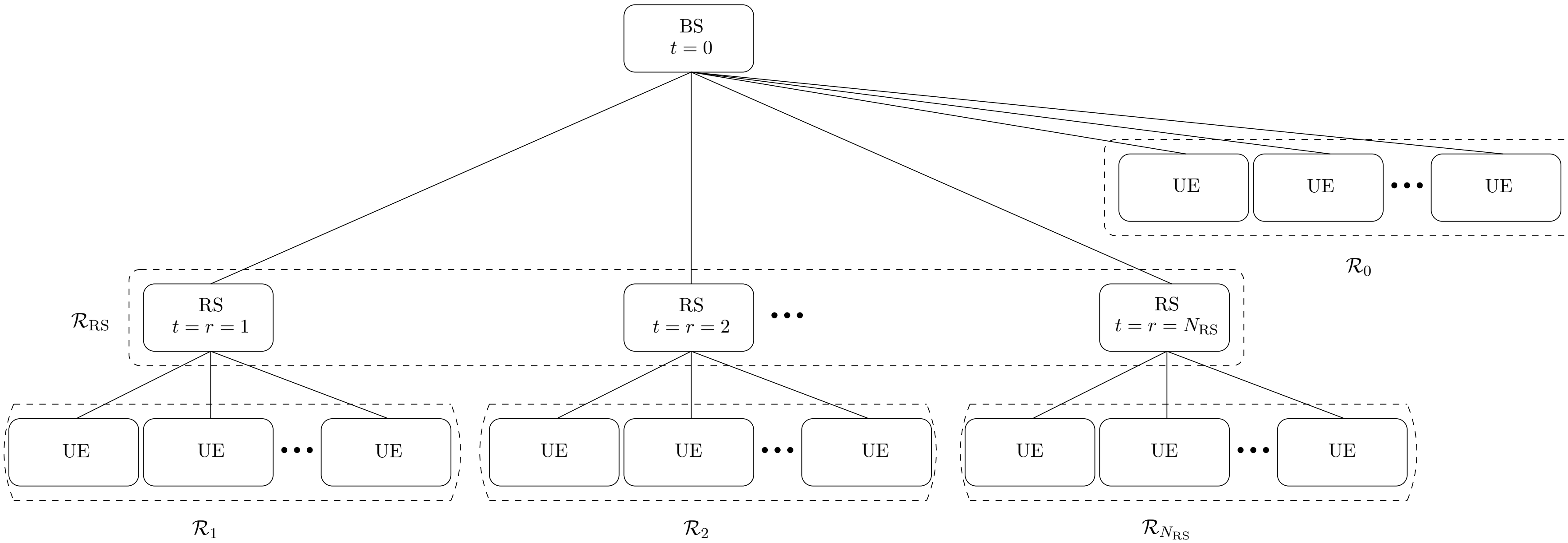

Figure 2.1. Topology of a cell in the relay network.

In the cell, the APs and UEs are indexed. The transmitting APs are represented by $t = 0, 1..., N_{\text{RS}}$. The BS is indicated by $t = 0$. The RSs are indexed by $1 \leq t \leq N_{\text{RS}}$. The receiving stations in a cell are the UEs and the RSs. A receiving station is represented by the index r, where $r = N_{\text{RS}} + 1, N_{\text{RS}} + 2..., N_{\text{RS}} + N_{\text{UE}}$. The receiving station is an RS for $1 \leq r \leq N_{\text{RS}}$ and a UE for $N_{\text{RS}} + 1 \leq r \leq N_{\text{UE}} + N_{\text{RS}}$. Note that the RS is both a transmitting and a receiving station in the downlink. A link is represented by a pair of indices (t, r), where t and r denote the transmitting AP and the receiving station, respectively. Throughout this thesis, the RSs and UEs which receive from an AP are represented by a set. The RSs are represented by $\mathcal{R}_{\text{RS}}$ and receive from the BS. If the indexes t or r represent an RS, this is written as $t \in \mathcal{R}_{\text{RS}}$ and $r \in \mathcal{R}_{\text{RS}}$, respectively. The UEs served by a transmitting AP t are represented by the set $\mathcal{R}_t$. The number of UEs served by an RS t is equal to the cardinality of the set $\mathcal{R}_t$, denoted by $|\mathcal{R}_t|$. The number of UEs served by the BS is $|\mathcal{R}_0|$ and the number of RSs is $|\mathcal{R}_{\text{RS}}| = N_{\text{RS}}$.

2.3 Medium Access of APs

In this section, the medium access is described for the downlink of the considered cell. Based on the description, a frame structure and various resource units are defined.

Since $N_{\text{RS}} + 1$ transmitting APs exist in the downlink of the cell, a medium access scheme is required enabling multiple transmissions in a frame. Two schemes are considered:

1. **Medium Access with Orthogonal Use of Time-Frequency Units**
   In this scheme, time-frequency units are used by an AP always exclusively, i.e, two APs of the cell do not transmit simultaneously on the same carrier frequency. Co-channel interference does not occur, but only noise exists and limits the user rates and sum rate. The scheme is called orthogonal medium access for the rest of this thesis. The orthogonal medium access models two kind of applications of the relay network. In the first one, the considered cell is the only one of the relay network and only a small number of APs exist such that a reuse of carrier frequencies occurs. In the second one, further cells exist, but these cells are neglected since APs of these cells are so far away that interfering signals transmitted by these APs are attenuated such that their co-channel interference vanishes in noise.

2. **Medium Access with Reuse of Time-Frequency Units**
   Time-frequency units are used by several APs of the cell in this scheme. Due

to the reuse, co-channel interference exists. The scenario is called reuse medium access for the rest of this thesis. The reuse medium access models the cell in a relay network in which co-channel interference limits the user rates and sum rate.

The transmissions of the APs are organized in frames. In order to illustrate the definition of resource units, Fig. 2.2 shows the frame structure. A frame has the duration of S slots. Slots are indexed by $s = 1, 2, ...S$. At least one set of successive slots is assigned to each AP. This set of slots is called subframe. The number of subframes is N_{SF} in a frame. In Fig. 2.2, the number of subframes is given by $N_{\mathrm{SF}} = 3$. The subframes are indexed by $n = 1, 2, ...N_{\mathrm{SF}}$. The duration of the subframe n is S_n slots and

$$\sum_{n=1}^{N_{\mathrm{SF}}} S_n = S \tag{2.1}$$

holds. The first subframe is reserved for the transmission of the BS. In orthogonal medium access, each subframe is allocated to an AP exclusively and $N_{\mathrm{SF}} \geq N_{\mathrm{RS}} + 1$ holds. In reuse medium access, APs of the same cell transmit simultaneously. Simultaneous transmission is realized by allocating subframe n to a group of APs. This group is called the reuse group $\mathcal{T}_n$. Finding reuse groups of the relay network depends on the specific scenario defining the expected data rate per AP, the environment and the position of the APs. APs causing a low interference to each other in their downlink transmissions are suitable for the same reuse group. For more details about finding reuse groups see [SWPI03, UMK06].

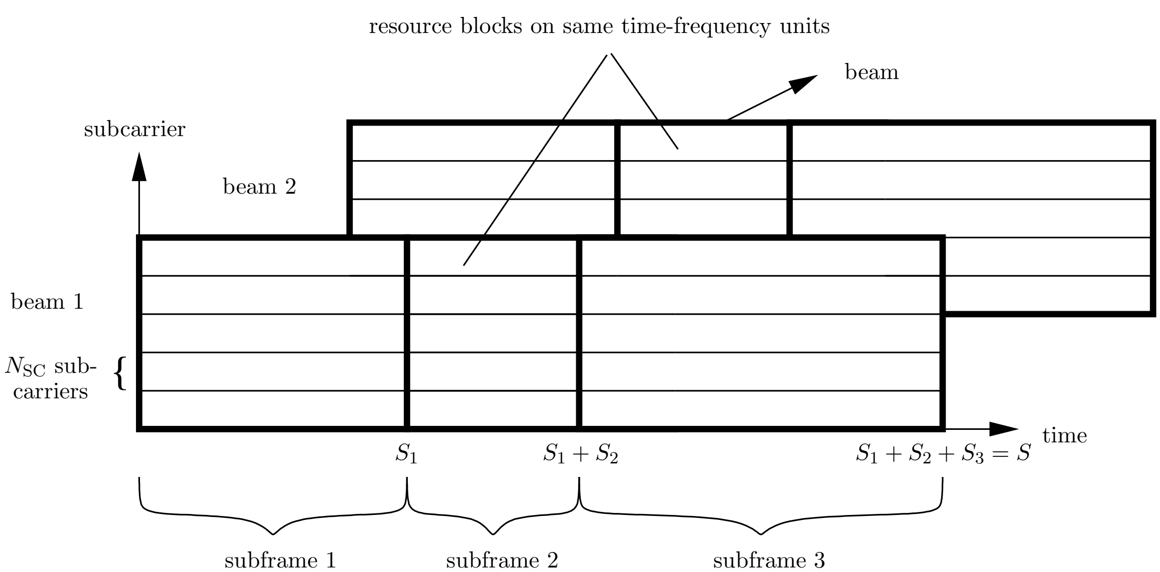

Figure 2.2. Frame structure with $N_{\mathrm{SF}=3}$ and $G_t = 2$ for all t.

Within its allocated subframe, an AP transmits to multiple UEs and to multiple RSs if the AP is a BS. The transmissions of the AP are multiplexed by OFDMA [NP00] combined with SDMA [PNG03]. A time frequency unit consists of N_{SC} adjacent subcarriers and has the duration of the subframe, given by S_n for subframe n. In frequency domain, F time-frequency units exist in subframe n. The time-frequency units in a subframe are indexed by $f = 1, 2, ...F$.

The number N_{SC} of subcarriers in a time frequency unit and the number S of slots in a frame are designed according to the properties of the wireless channel. In general, these properties are described by a time-variant transfer function [Pro95] modelling the channel depending on frequency and time. In order to enable efficient resource allocation algorithms as introduced in Chapter 4 and 5, two assumptions are made:

- The duration of a frame is sufficiently small: The coherence time [Pro95] gives the duration for which the time-variant transfer function can be considered to be time-invariant. Assuming that the frame duration is not larger than the coherence time, the transfer function can be considered as being time-invariant for one frame [Pro95]. This assumption implies that the UEs move with a low velocity.

- The bandwidth of a time-frequency unit is sufficient small: The coherence bandwidth [Pro95] gives the bandwidth over which the transfer function can be considered to be flat. Assuming that the bandwidth of a time-frequency unit is not larger than the coherence bandwidth, the time-invariant transfer function can be described by a complex scalar in a good approximation [Pro95].

If an AP t is equipped with more than one antenna, SDMA is applied. The AP t applies G_t beams and G_t data streams are multiplexed in space. Hence, a time-frequency unit is reused in space G_t times. In Fig. 2.2, it is assumed that $G_t = 2$. A resource block is defined by one beam and one time-frequency unit. The number of resource blocks is $K_t = G_t F$ for one subframe.

2.4 Modelling of the Transmit Unit and Receive Unit

In this section, the models of the transmit unit and of the receive unit are introduced in order to derive the restrictions to which the resource allocation problems introduced in Fig. 1.2 are subject. First the model of the transmit unit is presented, then the model of the receive unit. The focus of this section is on the transmit unit since the transmit unit consists of the entities enabling the resource allocation concepts introduced in Chapter 3.

A transmit unit is part of the BSs and RSs since they are the transmitting stations in downlink direction. The transmit unit is aware of CQI values. The transmit unit enables the design of grids of beams and the allocation of resource blocks, power and bits. Furthermore it performs the OFDM modulation. The restrictions concerning these possibilities are derived in the following. These restrictions are valid for each AP. For simplicity, the single AP is considered.

In order to enable SDMA, AP t is equipped with N_t antennas. For each time frequency unit, the antenna pattern is designed by choosing G_t beamforming vectors out of B_t ones. Each beamforming vector generates a beam. The G_t beams applied for the same time-frequency unit are called grid of beams. A beamforming vector is indexed by b, where $b = 1, 2, ...B_t$, and denoted by $\mathbf{m}_{t,b} \in \mathbb{C}^{N_t \times 1}$. Selecting G_t beams from B_t different beamforming vectors for each time-frequency unit and grouping beams in a grid of beams is called the design of grids of beams. The indicator variable $v_{t,b,k,n}$ is introduced in order to describe the design of grids of beams. The indicator variable is defined as

$$v_{t,b,k,n} = \begin{cases} 1 & \text{if beam } b \text{ of AP } t \text{ is allocated to resource block } k \text{ for subframe } n \\ 0 & \text{otherwise,} \end{cases} \tag{2.2}$$

where $k = 1, 2, ...K_t$. The design of grids of beams is subject to the constraint that G_t beams are used on a time-frequency unit f. This is written as

$$\sum_{k=(f-1)G_t+1}^{fG_t} \sum_{b=1}^{B_t} v_{t,b,k,n} = G_t, \ 1 \leq n \leq N_{\text{SF}} \text{ and } 1 \leq f \leq F. \tag{2.3}$$

Additionally, it must be considered that only one beam is selected for a resource block k. This is represented by

$$\sum_{b=1}^{B_t} v_{t,b,k,n} = 1, \ 1 \leq k \leq K_t \text{ and } 1 \leq n \leq N_{\text{SF}}. \tag{2.4}$$

The beamforming vector $\mathbf{m}_{t,k,n}$ applied to the resource block k is given by

$$\mathbf{m}_{t,k,n} = \sum_{b=1}^{B_t} v_{t,b,k,n} \mathbf{m}_{t,b}, \ 1 \leq k \leq K_t \text{ and } 1 \leq n \leq N_{\text{SF}}. \tag{2.5}$$

The transmit unit is able to apply multiple, different modulation and coding schemes to the resource blocks. A modulation symbol of a modulation and coding scheme carries ϵ data bits. The set of possible numbers of data bits is represented by $\mathcal{E}$. Assuming that a symbol duration is equivalent to a slot, the resource block k in subframe n is loaded with ϵS_n bits. Data bits are provided by a transmitting AP t to a receiving station with bit error probability $BEP_{t,r}$. A threshold expressed as an SINR value $\gamma_{\epsilon,r}$ must be exceeded to provide ϵS_n bits. The index r is used for the SINR value $\gamma_{\epsilon,r}$ since different bit error probability values may be demanded for receiving stations. The set $\mathcal{E}$ contains the value zero bits with smallest possible threshold of zero in order to define an allocation of bits if any other threshold cannot be exceeded.

The allocation of the resource block k to a link and the allocation of power and of bits to resource block k are described by a common assignment variable in order to model constraints including interdependencies between these problems. The assignment variable is defined by

$$u_{t,r,n,k,\epsilon} = \begin{cases} 1 & \text{if } \epsilon S_n \text{ bits are carried by resource block } k \\ & \text{allocated to link } (t,r) \text{ in subframe } n \\ 0 & \text{otherwise.} \end{cases} \tag{2.6}$$

Additionally, the auxiliary variable $\mathcal{R}$ is used in the following and represents the receiving stations of AP t. If the AP t is an RS, i.e., $t \neq 0$, then $\mathcal{R} = \mathcal{R}_t$. If the AP t is an BS, i.e., $t = 0$, then $\mathcal{R} = \mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}$. The assignment variable $u_{t,r,n,k,\epsilon}$ is linked with the allocation of power since the power can be allocated based on the known CQI such that the threshold $\gamma_{\epsilon,r}$ required for the transmission of ϵS_n bits is achieved. The power allocated to the resource block k and required to transmit ϵS_n bits is $p_{t,k}(\epsilon)$. The transmit power of AP t is limited by the value P_t. The allocation of power must fulfill the constraint

$$\sum_{r\in\mathcal{R}} \sum_{k=1}^{K_t} \sum_{\epsilon\in\mathcal{E}} p_{t,k}(\epsilon) u_{t,r,n,k,\epsilon} \leq P_t \tag{2.7}$$

if AP t transmits in subframe n.

The allocation of resource blocks is restriced by two constraints. Firstly, the data symbols addressed to the receiving stations in $\mathcal{R}$ are allocated to the resource blocks of subframe n. The AP t is not allowed to allocate resource blocks outside of the subframe n and the constraint

$$\sum_{r\in\mathcal{R}} \sum_{\epsilon\in\mathcal{E}} u_{t,r,n,k,\epsilon} = 0, \text{ if subframe } n \text{ is not allocated to the AP } t \tag{2.8}$$

$$\tag{2.9}$$

must be fullfilled for $1 \leq k \leq K_t$.

Secondly, only a limited number of resource blocks of the same time-frequency unit can be allocated to a receiving station r. In general, the number is limited by the number of antennas of the receiving station [PNG03]. In this thesis specific assumptions are made:

- A UE is alway equipped with a single antenna in this work. Hence, it cannot detect data carried by two resource blocks of the same time-frequency unit successfully.
- The number of beamforming vectors is limited in this thesis. The beams generated by the vectors differ in their directions. Since an RS is part of the infrastructure, the position of the RS is planned such that a LOS link is assumed between the transmitting BS and the receiving RS. Only one beam is directed in the LOS direction. Using the same beam twice for the same time-frequency unit causes a strong interference at the receiving RS. Using a beam not directed to a receiving RS leads to a low received signal power.

Because of these assumptions, up to one out of G_t resource blocks of the same time-frequency unit is allocated to a receiving station r in this thesis. This assumption yields to the constraint

$$\sum_{k=G_t(f-1)+1}^{G_t f} \sum_{\epsilon \in \mathcal{E}} u_{t,r,n,k,\epsilon} \leq 1, \text{ if subframe } n \text{ is allocated to AP } t. \tag{2.10}$$

and must be fulfilled for all $r \in \mathcal{R}$ and for all f.

The allocation of resource blocks and the allocation of bits is subject to the constraint that each resource block is allocated exactly once and that exactly one modulation and coding scheme is chosen for a resource block. This constraint is written as

$$\sum_{r \in \mathcal{R}} \sum_{\epsilon \in \mathcal{E}} u_{t,r,n,k,\epsilon} = 1, \ 1 \leq k \leq K_t. \tag{2.11}$$

if AP t transmits in subframe n.

Taking into account the solutions of the resource allocation problems, an OFDM modulation is applied for each antenna. The OFDM modulation performs an Inverse Fast Fourier Transformation (IFFT) and inserts a guard interval [NP00]. Subcarriers are assumed to be orthogonal and no inter-carrier interference occurs. The guard interval is designed such that inter-symbol interference is neglected.

The receive unit is part of the RSs and UEs since they are the receiving stations in the downlink direction. The receiving station is equipped with N_r antennas, where $N_r = 1$ if the station r is a UE. The receiving station applies an OFDM demodulator [PNG03] for each antenna branch. The OFDM demodulator removes the inserted guard interval and applies a Fast Fourier Transform (FFT) [NP00]. The bits carried by the resource blocks allocated to receiving station r are detected by the receive unit.

2.5 Signal Model

In this section, a signal model is introduced for an OFDMA transmission over the link between a transmitting AP t and a receiving station r. The link (t, r) corresponds to a direct connection or to one of the two links of a two-hop connection. The signal model is applied to derive the expression of the SINR value of the resource block k for orthogonal and reuse medium access. The expression of the SINR value is rather important since the SINR value must exceed the threshold γ_ϵ in order to transmit ϵ bits in a slot of a resource block.

According to Section 2.3, a resource block of subframe n consists of N_{SC} subcarriers and S_n slots, i.e., several OFDM symbols. For the sake of simplicity, a time-frequency unit is assumed to have the size of only one OFDM symbol and one subcarrier initially. This assumption is generalized in the end of this section. Since the signals of different subframes do not affect each other, the index n is not used in the signal model.

The variable $c_{t,k} \in \mathbb{C}$ represents the modulation symbols carried by resource block k. The modulation symbol $c_{t,k}$ is modelled by a zero-mean random variable and $c_{t,k}$ satisfies $\mathrm{E}\{c_{t,k}c_{t,k}^*\} = 1$, where $(\cdot)^*$ denotes the conjugate complex value and $\mathrm{E}\{\cdot\}$ the expectation operator. The alphabet of the modulation symbol is defined by the used modulation and coding scheme. The symbol is weighted by the power coefficient $\sqrt{p_{t,k}}$. The beamforming vector maps the symbol to the N_t antenna branches. The vector $\mathbf{m}_{t,k}$ is normalized such that $\mathbf{m}_{t,k}\mathbf{m}_{t,k}^{\mathrm{H}} = 1$, where $(\cdot)^{\mathrm{H}}$ denotes the conjugate transpose. The transmit signal $\mathbf{x}_{t,k} \in \mathbb{C}^{N_t \times 1}$ of the resource block k is given by

$$\mathbf{x}_{t,k} = \mathbf{m}_{t,k}\sqrt{p_{t,k}}c_{t,k}. \tag{2.12}$$

The transmit signal $\mathbf{x}_{t,k}$ consists of N_t elements indexed by $i = 1, 2, ..., N_t$. The i-th element of the transmit signal $\mathbf{x}_{t,k}$ is fed to the i-th antenna branch. At each antenna branch, the transmit signals of the resource blocks $(f-1)G_t + 1 \leq k \leq fG_t$ are summed up and fed to an OFDM modulator since the signals of G_t resource blocks are multiplexed on the same time-frequency unit.

The wireless channel between the OFDM modulator of transmit antenna i and the OFDM demodulator of receive antenna j, where $i = 1, 2, ..., N_t$ and $j = 1, 2, ..., N_r$, is modelled by the complex scalar $h_{t,r,f}^{(i,j)}$ for time-frequency unit f. The channel coefficient is affected by fast fading, slow fading and path loss [Par00]. Flat fading is assumed for each subcarrier and OFDM symbol of a time-frequency unit due to a proper design of the time-frequency units and of the frame as introduced in Section 2.3. The channel coefficients for all combinations of transmit antennas i and receive antennas j are represented by the matrix $\mathbf{H}_{t,r,f} \in \mathbb{C}^{N_r \times N_t}$.

The signal of resource block k received by station r is interfered by the signal of the other $G_t - 1$ resource blocks of the same time-frequency unit f. If reuse medium-access is assumed, the signal is also interfered by the APs of the same reuse group $\mathcal{T}_n$. Additionally, additive white Gaussian noise represented by the noise vector $\mathbf{z}_{r,k} \in \mathbb{C}^{N_r \times 1}$ disturbs the receiption. The elements of $\mathbf{z}_{r,k}$ are zero-mean circulant symmetric complex Gaussian random variables [Pap84] with a variance of Z_0. The receive signal $\mathbf{y}_{r,k} \in \mathbb{C}^{N_r \times 1}$ is given by

$$\mathbf{y}_{r,k} = \mathbf{H}_{t,r,f}\mathbf{x}_{t,k} + \sum_{\substack{k'=(f-1)G_t+1 \\ k' \neq k}}^{fG_t} \mathbf{H}_{t,r,f}\mathbf{x}_{t,k'} + \sum_{\substack{t' \in \mathcal{T}_n \\ t' \neq t}} \sum_{k'=(f-1)G_{t'}+1}^{fG_{t'}} \mathbf{H}_{t',r,f}\mathbf{x}_{t',k'} + \mathbf{z}_{r,k}. \quad (2.13)$$

The four terms correspond to the signal, inter-beam interference, co-channel interference and noise part, respectively. For orthogonal medium access, the co-channel interference part is zero. A linear receive filter is assumed at the receive unit in order to compensate phase shifts and amplitude variations caused by the wireless channel. The filter is denoted by $\mathbf{d}_{r,k} \in \mathbb{C}^{N_r \times 1}$. The filtered signal $\hat{y}_{r,k}$ is given by

$$\hat{y}_{r,k} = \mathbf{d}_{r,k}^{\mathrm{T}} \mathbf{y}_{r,k}, \quad (2.14)$$

where $(\cdot)^{\mathrm{T}}$ denotes the transpose. In order to derive the expression of the SINR, the received signal power, the power of the inter-beam interference, the power of the co-channel interference and the noise power must be determined.

For a deterministic beamforming vector, filter and wireless channel and for random modulation symbols, the received signal power is

$$\begin{aligned} P_{r,k} &= \mathrm{E}\{(\mathbf{d}_{r,k}^{\mathrm{T}}\mathbf{H}_{t,r,f}\mathbf{x}_{t,k})^{\mathrm{H}}\mathbf{d}_{r,k}^{\mathrm{T}}\mathbf{H}_{t,r,f}\mathbf{x}_{t,k}\} \\ &= p_{t,k}\mathbf{m}_{t,k}^{\mathrm{H}}\mathbf{H}_{t,r,f}^{\mathrm{H}}(\mathbf{d}_{r,k}^{\mathrm{T}})^{\mathrm{H}}\mathbf{d}_{r,k}^{\mathrm{T}}\mathbf{H}_{t,r,f}\mathbf{m}_{t,k}. \end{aligned} \quad (2.15)$$

In order to simplify (2.15), the channel gain is defined as

$$\alpha_{t,r,k}^2 = \mathbf{m}_{t,k}^{\mathrm{H}}\mathbf{H}_{t,r,f}^{\mathrm{H}}(\mathbf{d}_{r,k}^{\mathrm{T}})^{\mathrm{H}}\mathbf{d}_{r,k}^{\mathrm{T}}\mathbf{H}_{t,r,f}\mathbf{m}_{t,k} \quad (2.16)$$

and the received signal power is rewritten as

$$P_{r,k} = p_{t,k}\alpha^2_{t,r,k}. \tag{2.17}$$

Note that $\alpha^2_{t,r,k}$ is a function of the assignment variable $v_{t,b,k,n}$ according to (2.5). The power $I_{\mathrm{IBI},r,k}$ of the inter-beam interference caused by the resource blocks k', where $(f-1)G_t+1 \leq k \leq fG_t$ and $k' \neq k$, is determined by

$$\begin{aligned} I_{\mathrm{IBI},r,k} &= \sum_{\substack{k'=(f-1)G_t+1 \\ k'\neq k}}^{fG_t} \mathrm{E}\{(\mathbf{d}^{\mathrm{T}}_{r,k}\mathbf{H}_{t,r,f}\mathbf{x}_{t,k'})^{\mathrm{H}}\mathbf{d}^{\mathrm{T}}_{r,k}\mathbf{H}_{t,r,f}\mathbf{x}_{t,k'}\} \\ &= \sum_{\substack{k'=(f-1)G_t+1 \\ k'\neq k}}^{fG_t} p_{t,k'}\alpha^2_{t,r,k'}. \end{aligned} \tag{2.18}$$

The power $I_{\mathrm{CCI},r,k}$ of the co-channel interference is calculated by

$$\begin{aligned} I_{\mathrm{CCI},r,k} &= \sum_{\substack{t'\in\mathcal{T}_n \\ t'\neq t}} \sum_{k=(f-1)G_{t'}+1}^{fG_{t'}} p_{t',k}\mathrm{E}\{(\mathbf{d}^{\mathrm{T}}_{r,k}\mathbf{H}_{t',r,f}\mathbf{x}_{t',k})^{\mathrm{H}}\mathbf{d}^{\mathrm{T}}_{r,k}\mathbf{H}_{t',r,f}\mathbf{x}_{t',k}\} \\ &= \sum_{\substack{t'\in\mathcal{T}_n \\ t'\neq t}} \sum_{k=(f-1)G_{t'}+1}^{fG_{t'}} p_{t',k'}\alpha^2_{t',r,k'}. \end{aligned} \tag{2.19}$$

The noise power is given by

$$Z_{r,k} = E\{(\mathbf{d}^{\mathrm{T}}_{r,k}\mathbf{z}_{r,k})^{\mathrm{H}}\mathbf{d}^{\mathrm{T}}_{r,k}\mathbf{z}_{r,k}\} = ||\mathbf{d}_{r,k}||^2 N_r Z_0, \tag{2.20}$$

where $||\cdot||$ denotes the absolute value. The SINR of the resource block k is

$$\gamma_{r,k} = \frac{P_{r,k}}{Z_{r,k} + I_{\mathrm{IBI},r,k} + I_{\mathrm{CCI},r,k}}. \tag{2.21}$$

So far, it is assumed that the time-frequency unit f and the resource block k consist of one OFDM symbol and one subcarrier. In practice, several OFDM symbols and subcarriers are grouped in time and frequency domain, respectively, to a time-frequency unit and resource block [3GP06, IEE04] in order to reduce signaling. The SINR value $\gamma_{r,k}$ of (2.21) is valid for a resource block if the following constraints are fulfilled:

- The channel matrix $\mathbf{H}_{t,r,f}$ is valid for all OFDM symbols in resource block k.
- The same modulation and coding scheme is applied for all OFDM symbols in resource block k.
- The allocation of power $p_{t,k}$ is valid for all OFDM symbols in resource block k.

- The same beamforming vector $\mathbf{m}_{t,k}$ is used for all OFDM symbols in resource block k.

The first constraint is ensured in this thesis since a proper design of the time-frequency unit f taking into account the coherence time and bandwidth is assumed, cf. Section 2.3. The last three constraint are fulfilled in this thesis since the allocation of power and bits and the design of grids of beams is applied per resource block.

2.6 Performance Measures

In order to describe the performance of a link, a connection and the considered cell, measures are required. Based on the models introduced in this chapter, the data rate of the link (t, r), the user rate of a direct connection and of a two-hop connection as well as the sum rate and the average sum rate are defined in this sections. These definitions are used to formulate the objective of the resource allocation problems in Chapter 3 and to evaluate the performance of the algorithms proposed in this thesis.

The data rate of the link (t, r) is defined as the number of bits allocated by the transmitting AP t to the receiving RS or UE r and is related to the duration of a frame. The data rate of the link (t, r) is given in bits per frame and results from counting all bits allocated to the link (t, r) in a frame. The data rate is represented by

$$R_{t,r} = \sum_{n=1}^{N_{\mathrm{SF}}} \sum_{k=1}^{K_t} \sum_{\epsilon \in \mathcal{E}} u_{t,r,n,k,\epsilon} \epsilon S_n. \tag{2.22}$$

Note that the data rate is offered with a bit error probability $BEP_{0,r}$ according to the threshold $\gamma_{\epsilon,r}$. The definition of the data rate of the link (t, r) is used to define the user rates between the BS and the UEs. If the connection between the BS $t = 0$ and the UE r is a direct connection, i.e., $r \in \mathcal{R}_0$, the user rate is given by the data rate $R_{0,r}$ in (2.22).

If the connection between the BS $t = 0$ and the UE r is a two-hop connection via the RS $r' = t'$, where $r', t' \in \mathcal{R}_{\mathrm{RS}}$, the bits are transmitted first to RS r'. The data rate of the information sent to RS r' and finally addressed to UE r is denoted as $R_{0,r',r}$. The user rate is given by the minimum of the data rate $R_{0,r',r}$ and the data rate of the link (t', r). The assumption is made that bits transmitted to RS r' and not forwarded to the addressed UE r in the same frame are lost. This assumption avoids that delays of more than one frame are introduced by the radio transmission. The user rate of a two-hop connection is given by

$$R_{0,r} = \min\{R_{0,r',r}; R_{t',r}\}. \tag{2.23}$$

If the data rate of the BS-to-RS link is larger than the data rate of the RS-to-UE link,i.e.,

$$R_{0,r',r} \geq R_{t',r}, \tag{2.24}$$

the user rate is given by the data rate $R_{t',r}$ since the RS cannot forward more than the RS-to-UE link allows. If the data rate of the BS-to-RS link is smaller than the data rate of the RS-to-UE link, the user rate is given by the data rate $R_{0,r',r}$ since the RS cannot forward more than it receives from the BS.

The bit error probability of a two-hop connection is determined using an approximation introduced in [FT07]. The approximation neglects that a bit error which occurs on the BS-to-RS link is corrected on the RS-to-UE link at random [FT07]. The bit error probability $BEP_{0,r}$ on a two-hop connection is approximated by

$$BEP_{0,r} \approx 1 - (1 - BEP_{0,r'})(1 - BEP_{t',r}). \tag{2.25}$$

The sum rate $R_{\sum}$ of the cell is the sum of all user rates

$$R_{\sum} = \sum_{r \in \bigcup_{t=0}^{N_{\mathrm{RS}}} \mathcal{R}_t} R_{t,r}. \tag{2.26}$$

The sum rate $R_{\sum}$ varies from frame to frame since the wireless channel is time varying. In order to provide a generalized performance measure, the average sum rate is defined as the arithmetic mean of the sum rate measured over multiple frames.

Chapter 3
Problem Formulations and Resource Allocation Concepts

3.1 Introduction

The resource allocation problems shown in Fig. 1.2 affect each other. A straightforward approach is to formulate a joint problem and to solve it. In Section 3.2, two joint formulations are presented: one with the objective of maximizing the minimum user rate and one with the objective of maximizing the sum rate. Their solutions lead to optimum resource allocations in a cell of the relay network, but finding their solutions is not applicable under realistic assumptions. As shown in the end of Section 3.2, the required computational complexity and the required amount of signalling between the APs are too large. Therefore, solving the joint formulations is not continued. Instead, two concepts called distributed concept for orthogonal medium access and distributed concept for reuse medium access are proposed in Section 3.3. Each concept is applicable to both objectives. The concepts differ since a concept is defined for each medium access scheme considered in Section 2.3. Each concept decomposes the joint problems in subproblems, defines the successive order in which the subproblems are solved and defines which subproblem is solved by the BS and which one by the RSs. Section 3.4 summarizes the main contributions of this chapter.

3.2 Joint Formulation of Resource Allocation Problems

In this section, two joint formulations of the resource allocation problems illustrated in Fig. 1.2 are motivated and given. Two objectives are considered, hence, two formulations are presented. The first one aims at maximizing the minimum user rate in a cell of the relay network. The second one aims at maximizing the sum rate. For both objectives, a general formulation is given valid for the orthogonal medium access and reuse medium access. It is shown that it is not applicable to solve both problems in an optimum way in a relay network under realistic assumptions since the required computational complexity and the required amount of signalling between the APs are too large.

The joint formulation is motivated as follows. Resource block k in subframe n is considered. Subframe n has the length of S_n slots. The resource block carriers ϵS_n

bits. The success of the transmission requires that the SINR $\gamma_{r,k}$ fulfills

$$\gamma_{r,k} \geq \gamma_{\epsilon,r}. \tag{3.1}$$

The design of grids of beams and the allocation of resource blocks, power and bits must ensure that the received signal power $P_{r,k,n}$ per resource block fulfills

$$P_{r,k,n} = \alpha^2_{t,r,k,n} p_{t,k,n} \geq \gamma_{\epsilon,r}(Z_{r,k,n} + I_{\mathrm{IBI},r,k,n} + I_{\mathrm{CCI},r,k,n}), \tag{3.2}$$

where (2.17) and (2.21) are applied to (3.1). According to Section 2.5 the power of the co-channel interference is zero for the orthogonal medium access, i.e., $I_{\mathrm{CCI},r,k} = 0$. For the reuse medium access, the power of the co-channel interference is larger than zero, i.e., $I_{\mathrm{CCI},r,k} > 0$. Hence, equation (3.2) is valid for the orthogonal medium access and the reuse medium access and the following considerations are valid for both medium access schemes unless otherwise stated.

The power of each transmitting AP t is limited according to (2.7) and must be utilized efficiently. It is sufficient for the transmission of ϵS_n bits to fulfill (3.2) with equality. Then, the allocated power $p_{t,k,n}$ is

$$p_{t,k,n} = \frac{\gamma_{\epsilon,r}(Z_{r,k,n} + I_{\mathrm{IBI},r,k,n} + I_{\mathrm{CCI},r,k,n})}{\alpha^2_{t,r,k,n}}, \tag{3.3}$$

if ϵS_n bits are transmitted. Equation (3.3) and its derivation show that the successful transmission of ϵS_n bits is affected by

- the design of grids of beams since the beamforming vector applied for resource block k in subframe n affects the channel gain value $\alpha^2_{t,r,k,n}$ according to (2.16) and the channel gain values determine the power $I_{\mathrm{IBI},r,k,n}$ of the inter-beam interference and the power $I_{\mathrm{CCI},r,k}$ of the co-channel interference according to (2.18) and (2.19), respectively,

- the allocation of resource blocks since the channel gain value $\alpha^2_{t,r,k,n}$, the SINR threshold $\gamma_{\epsilon,r}$, the noise power $Z_{r,k,n}$, the power $I_{\mathrm{CCI},r,k,n}$ of the co-channel interference and the power $I_{\mathrm{IBI},r,k,n}$ of the inter-beam interference depend on the receiver as denoted by the index r,

- the allocation of power as represented by the variable $p_{t,k,n}$,

- the allocation of bits since the transmission of ϵS_n bits requires to meet the threshold $\gamma_{\epsilon,r}$,

- the allocation of slots since it determines the number S_n of slots in the resource block.

Since the resource allocation problems illustrated in Fig. 1.2 affect each other, joint formulations are given for the two objectives introduced in Section 1.2. Formulations as integer programs given by the objective and its constraints are chosen since the number of beamforming vectors, the number of time-frequency units, the number of modulation and coding schemes and the number of slots are integers. Even the power, which is actually a continuous quantity, can only reach an integer number of possible values since it is sufficient to meet (3.3) for each t, k and n with equality. Once the satisfying bit error probability is ensured, allocating more power to a resource block is only beneficial if a modulation and coding scheme can be applied carrying more bits.

The first problem is defined as maximizing the minimum user rate in the cell. The problem is called (P1). The variables $u_{t,r,k,n,\epsilon}$, $v_{t,r,k,n}$ and S_n introduced in (2.6), (2.4) and (2.1) are used to describe the design of grids of beams, the allocation of resource blocks, power and bits and the allocation of slots. The minimum is searched over the union $\bigcup_{t=0}^{N_{\mathrm{RS}}} \mathcal{R}_t$ corresponding to all BS-to-UE and RS-to-UE links. Using the definition of the user rate $R_{0,r}$ in Section 2.6, the objective of (P1) is written as

$$\max_{\substack{u_{t,r,k,n,\epsilon}\\ v_{t,r,k,n}\\ S_n}} \min_{r \in \bigcup_{t=0}^{N_{\mathrm{RS}}} \mathcal{R}_t} R_{0,r} = \max_{\substack{u_{t,r,k,n,\epsilon}\\ v_{t,r,k,n}\\ S_n}} \min_{r \in \bigcup_{t=0}^{N_{\mathrm{RS}}} \mathcal{R}_t} \sum_{n=1}^{N_{\mathrm{SF}}} \sum_{k=1}^{K_t} \sum_{\epsilon \in \mathcal{E}} u_{t,r,k,n,\epsilon} \epsilon S_n. \tag{3.4}$$

The objective is subject to several constraints ordered in five classes. The classes are related to the definition of the variables, the three resource allocation problems illustrated in Fig. 1.2 and the data rate. Multiple constraints are already discussed in Chapter 2. For simplicity, the references to these constraints are only listed here. The constraints are:

1 Variable constraints:

   1.1 $u_{t,r,k,n,\epsilon} \in \{0, 1\}$

   1.2 $v_{t,r,k,n} \in \{0, 1\}$

   1.3 $S_{n,\epsilon} \in \{0, 1, ..., S\}$

2 Constraints for the design of grids of beams:

   2.1 A grid of beam is designed for each time-frequency unit and consist of G_t beams according to (2.3).

   2.2 Exactly one beam is used for a resource block according to (2.4).

3 Constraints for the allocation of resource blocks, power and bits:

   3.1 The AP t is not allowed to allocate resource blocks outside of the subframe n according to (2.8).

3.2 At most one resource block of the G_t resource blocks of the same time-frequency unit is allocated to receiver r according to (2.10).

3.3 The power allocated to a resource block is given by (3.3). The power of each transmitting access point t is limited according to (2.7).

3.4 Each resource block is allocated exclusively to one receiver and combined with a modulation and coding scheme according to (2.11).

4 Constraint for the allocation of slots:

4.1 The frame is split in subframes. The sum of the lengths of the subframes is limited by the frame size according to (2.1).

5 Data rate constraints:

5.1 The data rate of the second hop must be provided on the first hop according to (2.24).

The second problem is defined as maximizing the sum rate and called (P2). Its formulation uses the same variables as in the problem (P1). Using the definition of the sum rate $R_{\sum}$ in equation (2.26), the objective of (P2) is written as

$$\max_{\substack{u_{t,r,k,n,\epsilon} \\ v_{t,r,k,n} \\ S_n}} R_{\sum} = \max_{\substack{u_{t,r,k,n,\epsilon} \\ v_{t,r,k,n} \\ S_n}} \sum_{r \in \bigcup_{t=0}^{N_{\mathrm{RS}}} \mathcal{R}_t} \sum_{n=1}^{N_{\mathrm{SF}}} \sum_{k=1}^{K_t} \sum_{\epsilon \in \mathcal{E}} u_{t,r,k,n,\epsilon} \epsilon S_n. \tag{3.5}$$

The objective is subject to the same constraints as problem (P1). Additionally, the constraint 5.2 is defined:

5.2 Each UE is allocated at least the minimum user rate $R_{\mathrm{min},0,r}$ written as $R_{\mathrm{min},0,r} \leq R_{t,r}$ for each UE r.

The solutions of problems (P1) and (P2) are restricted by two prerequisites. At first, information about the wireless channels in the considered cell is required. This information is only gained if a large signalling is accepted between the RSs and the BS. Secondly, a huge computational complexity is required.

The large signalling as stated in the first prerequisite is derived as follows. In order to solve both problem, information about the noise power and CSI is required. At first, a single link (t, r) is considered. In a Time Division Duplex (TDD) system, the CSI is estimated by the AP t when AP t is receiving the uplink transmission send by station

r. In a Frequency Division (FDD) system, it is estimated by the receiving station r during the downlink transmission and fed back to the AP t. The noise power is always fed back from the receiving station to the AP. If multiple links are considered, the AP t gathers information about the noise power and CSI of all served links. If the problems (P1) and (P2) shall be solved, the information about the noise power and CSI will be required for each link (t, r) in the cell. For the reuse medium access, the channel gain values are additionally required for each channel between the receiver $r \in \mathcal{R}_t$ and the interfering APs, included in the set $\mathcal{T}_n$ for subframe n. Additionally, the power allocated to time-frequency units by the interfering APs must be known. Since the total information is distributed over the APs in the cell, all the information must be feed to a central unit (e.g., the BS) per frame in order to solve problem (P1) and (P2) once per frame. Feeding all these values once in the duration of a frame - typically in a few milliseconds [IEE04,3GP06] - to a central unit causes an intolerable feedback overhead.

Even if it is assumed that all the required information is given, both problems will require a huge computational complexity as as stated in the second prerequisite and motivated in the following. The problem (P1) and (P2) are nonlinear integer programs. Both problems can be solved by testing all possible combinations of the variables $u_{t,r,k,n,\epsilon}$, $v_{t,r,k,n}$ and S_n. This brute-force solution is called exhaustive search algorithm. However, the complexity of the exhaustive search algorithm increases exponentially with the input parameters, i.e., the number of possible beamforming vectors, the number of resource blocks, the number of modulation and coding schemes, the number of users in the cell and the number of slots in a frame. An exhaustive search is impossible for realistic assumptions about these parameters. A solution with an acceptable computational complexity is still an open research topic. A possibility is the application of the branch-and-bound algorithm to the nonlinear integer programs [Ken88]. But its complexity is still growing exponentially with the number of input parameters [DZC99] in the worst case and therefore, it is not suitable for a real-time application.

Both prerequisites prevent to find the optimal solution of problems (P1) and (P2) under realistic assumptions.

3.3 Decomposition of the Resource Allocation Problems

Concepts leading to an applicable but suboptimum solution for the problems (P1) and (P2) for the orthogonal and reuse medium access are required. Two concepts are proposed in this section called distributed concept for orthogonal medium access and distributed concept for reuse medium access. The concepts differ in the considered medium access scheme. The former one is only applicable to the orthogonal medium access since co-channel interference is not considered. The latter one is designed for the reuse medium access, but it can also be applied to the orthogonal medium access as shown in the course of this section. Since both problems (P1) and (P2) differ only in the objective and in the constraint 5.2 of problem (P2), a common concept is proposed for one medium access scheme. The different objectives are considered if a concept is applied to one of the problems (P1) and (P2). In the following, the motivation of the concepts is given, then the concepts are presented.

Based on the considerations of the previous section, the concepts solving problem (P1) or (P2) must take into account limitations concerning:

1. the computational complexity,
2. the amount of signalling between BS and RSs of the cell.

Based on the consideration of Section 1.3, the concepts must take into account the limitation that

3. CSI is hard to obtain [VTL02].

These limits are considered in both concepts as follows:

1. The computational complexity is reduced by decomposing the problems (P1) and (P2) in subproblems. These subproblems are related to the design of grids of beams, the allocation of resource blocks, power and bits and the allocation of slots. The definitions of these problems differ for the objectives aiming at maximzing the minimum user rate and maximizing the sum rate. The interdependencies occurring between the subproblems and illustrated in Fig. 1.2 is considered in the definition of the subproblems.

2. The signalling between BS and RSs is kept low by solving the allocation of resource blocks, power and bits partly by the BS and partly by the RSs and not only by the BS as a central unit. Solving this subproblem by the RS t makes it possible that the RS t signals only the number of allocated bits per slot about each link (t, r), where $r \in \mathcal{R}_t$, to the BS. Information about the resource blocks is only required at RS t for each link (t, r), where $r \in \mathcal{R}_t$ and at the BS for each link $(0, r)$, where $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}$. CQI or CSI related to the resource blocks needs not to be sent to the BS for each RS-to-UE link as required for an optimal

solution of problems (P1) and (P2) as found by the exhaustive search algorithm. Furthermore, parts of the computational load is distributed over all APs in the cell and several operations are performed in parallel. Note that the other sub-problems are solved by the BS. The design of grids of beams is solved by the BS in order to adapt the usage of the beams to the distribution of the UEs in the cell. The allocation of slots is also solved by the BS since all APs of the cell compete for the slots.

3. In order to gain information about the wireless channel, the concepts use CQI values and not CSI values as motivated by the concept of opportunistic beamforming [VTL02]. In the concept of opportunistic beamforming, the grids of beams of a single BS are designed firstly, then CQI values defined by SINR values are collected via a feedback channel from the receivers and finally the resource blocks are allocated to links. An extension of this concept to multiple BSs and cells of a conventional networks is proposed in [GRB06b, HR08]. The advantage of the concept is that only CQI values must be estimated and, if necessary, fed back by the receivers and not the complex valued CSI. This reduces the amount of feedback and makes the concept rather robust against estimation and feedback errors [VTL02, GRB06b].

The RS t requires the CQI values for each link (t, r), where $r \in \mathcal{R}_t$ and the BS requires the CQI values for each link $(0, r)$, where $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}$ in order to solve the allocation of resource blocks, power and bits. The question arises by which parameters the CQI values are defined and how the CQI values are gained. Since co-channel interference occurs in reuse medium access, but not in orthogonal medium access, two different types of CQI values are defined for the two medium access schemes in this thesis.

Firstly, only the orthogonal medium access is considered. The first type of CQI values are the channel gain values $\alpha_{t,r,k,n}$ for each resource block k of subframe n as defined in (2.16) and the noise power $Z_{r,k}$ defined for each link (t, r) in the cell. The CQI values are gained by the following method. At first, the design of grids of beams is solved without any CSI as in [VTL02, GRB06b]. Then, the APs transmit pilots on all resource blocks in a pilot phase. The APs transmit their pilots successively in time according to their order in the frame structure. Each receiving RS and UE estimates the CQI values describing the link to its transmitting AP. In a following feedback phase, the CQI values are sent from the receiving RSs and UEs to the BS and RS, respectively. Each RS t of the cell is sent the CQI values of the links (t, r) where $r \in \mathcal{R}_t$. The BS is sent the CQI values of the links $(0, r)$, where $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}$. After the feedback phase, the allocation of resource blocks, power, bits and slots is solved. Since co-channel interference does not occur, each AP gains the CQI values autonomously.

Now only the reuse medium access is considered. The second type of CQI values are SINR values. Collecting channel gain values is too complex since the channel gain values between all APs of the same reuse group and the receiving UE and RS are required to solve the allocation of resource blocks, power and bits. The SINR values can be obtained easier than channel gain values as proposed in [GRB06b]: The subproblem of the design of grids of beams is solved and a power value is allocated to each resource block. Since the allocation of resource blocks to links is not known yet, the same power is allocated to each resource block. In the following pilot phase, all APs of the same reuse group transmit their pilots simultaneously taking into account the pre-allocated beamforming vectors and power values. The pilot phase must be coordinated across all APs of the reuse groups. A pilot phase is required for each reuse group. A receiving UE r, where $r \in \mathcal{R}_t$ estimates the SINR values of the link (t, r) for all the K_t resource blocks of a subframe. A receiving RS estimates the SINR values of the link to the BS. In the following feedback phase, these SINR values are fed back by to the APs. During the transmission in the frame it is important that the same beamforming vectors and the same power is applied to the resource blocks as in the pilot phase in order to keep the SINR values valid. Hence, the allocation of power must be solved before the SINR values are available at the APs. Even the allocation of slots must be solved before the pilot phase since the reuse group must not change afterwards. The SINR values fed back to the APs are the second type of CQI values. Note that this type of CQI values can be also gained by the same method if orthogonal medium access is used. The reason is that the orthogonal medium access can be seen as a reuse medium access in which each reuse group consists of only one AP. However, using the second type of CQI values leads to a performance degradation compared to the first type of CQI values since the allocation of power and allocation of slots is solved without CQI knowledge.

Since two types of CQI values and two methods how to gain these CQI values are defined, two resource allocation concepts are introduced in the following based on the motivation given so far.

The first concept is called distributed concept for orthogonal medium access. The CQI values are the channel gain values and noise power values. A flow chart of the concept is depicted in Fig. 3.1. The flow chart shows the successive order of the subproblems which are solved. On the left hand side of Fig. 3.1, the required information needed to solve the subproblems are given if existing. On the right hand side, the APs solving

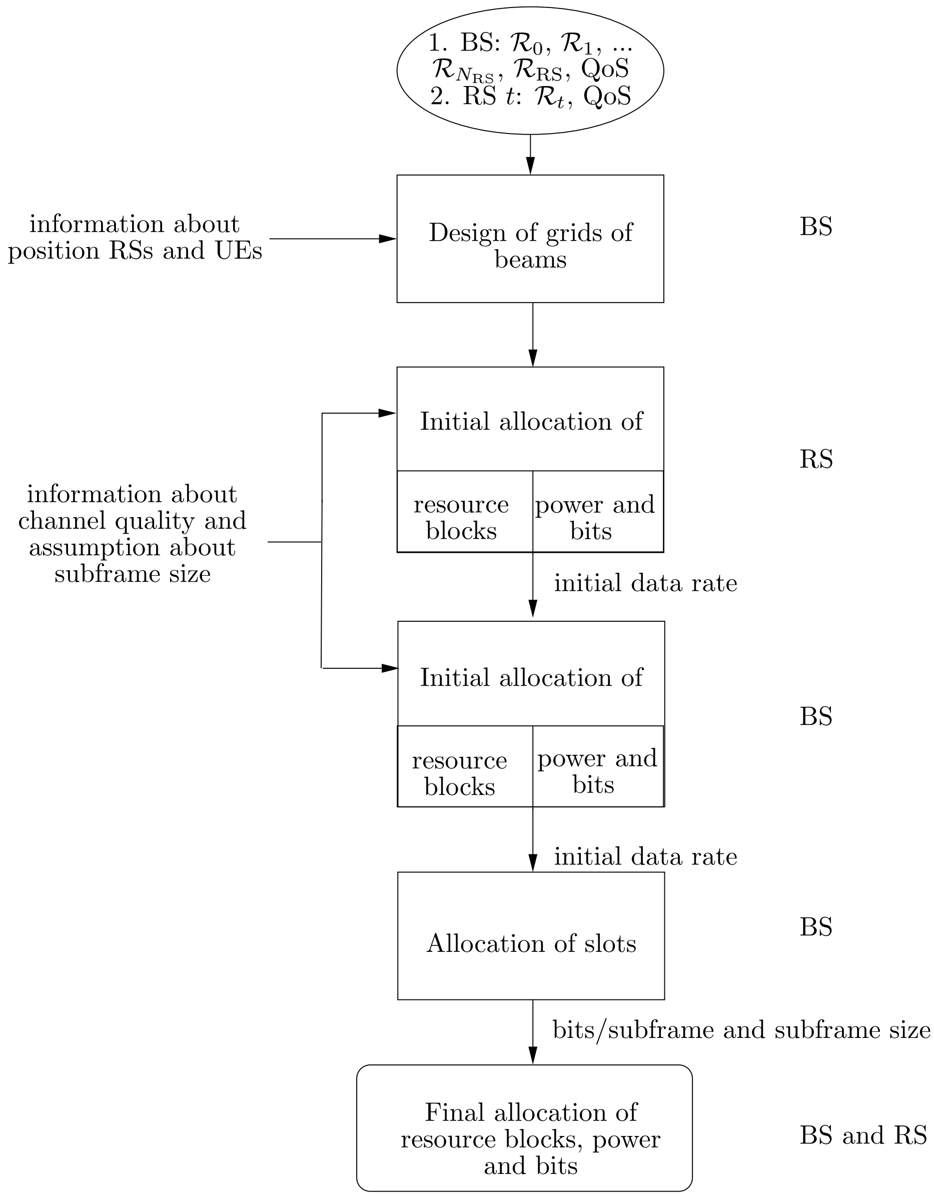

Figure 3.1. Flow chart of the distributed concept for orthogonal medium access.

the subproblems are stated. The distributed concept for orthogonal medium access is defined as follows:

1. Initialization
   Initially, the BS knows all existing connections and links in the cell, i.e., the sets $\mathcal{R}_0$, $\mathcal{R}_1$, ... $\mathcal{R}_{N_{\mathrm{RS}}}$ and $\mathcal{R}_{\mathrm{RS}}$. For each link, the bit error probability $BEP_{t,r}$, which must be fulfilled, is known. If the objective is the maximization of the sum rate, the minimum data rate $R_{\min,t,r}$ is also known for each link. Both parameters together are called QoS in Fig. 3.1. The RS $t \in \mathcal{R}_{\mathrm{RS}}$ knows the receiving stations served by itself and given by $\mathcal{R}_t$. Furthermore, it knows the bit error probability

$BEP_{t,r}$ and if existing, the minimum data rate $R_{\min,t,r}$ for each link (t, r) where $r \in \mathcal{R}_t$.

2. Design of grids of beams
   The design of grids of beams is solved by the BS. In addition to the information known by the initialization, the BS gets and uses information about the position of the RSs and the position of the UEs in the cell. Since the RSs are fixed, the exact position of the RSs are assumed to be known at the BS. The exact position of the UEs are not easy to obtain since the UEs are nomadic or mobile. Here, it is assumed that the probability density function $\rho(x, y)$ is given which defines the probability that the UE r is at the position (x, y) within the cell. The solution of the subproblem of design of grids of beams defines the beamforming vector used for each resource block. It enables to determine the CQI for each resource block and each receiving station as stated in the definition of the types of CQI values.

3. Initial allocation of resource blocks, power and bits by the RS
   When the RS t knows the CQI values of all links (t, r) where $r \in \mathcal{R}_t$, the RS t determines an initial allocation of resource blocks, power and bits for the links (t, r) where $r \in \mathcal{R}_t$. In order to reduce the computational complexity, the subproblem is divided further in the allocation of resource blocks and the allocation of power and bits. Since the size of the subframe in which RS t transmits is not known a priori, the subframe size is assumed to be one slot. For one slot, the allocation of resource blocks and the allocation of power and bits are considered. Both subproblems are solved in successive order to reduce the complexity. As the result, the RS knows the bits per slot which could be transmitted from the RS to each UE. This quantity is called initial data rate and is signaled to the BS for each UE $r \in \mathcal{R}_t$.

4. Initial allocation of resource blocks, power and bits by the BS
   When the BS knows the CQI values for its links $(0, r)$ where $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\text{RS}}\}$ and the initial data rate values of the RS-to-UE links, the BS makes an initial allocation of resource blocks, power and bits for its links. Again, the subproblem is further divided into the allocation of resource blocks and allocation of power and bits. The size of the subframe is assumed to be one since the size is not known yet. As the result, the BS knows the initial data rate values of all links in the cell.

5. Allocation of slots
   Based on the initial data rate values, the BS allocates slots to the APs, i.e., determines the sizes of the subframes. Since a subframe size of one slot is assumed so far, the data rate of a link is a linear function of the number of slots allocated

to the subframe in which the link is served. As the result, the BS sends the subframe size and initial data rate values of the BS-to-RS links to the RSs. The allocation of slots is solved after the initial allocation of resource blocks, power and bits in order to allocate slots taking into account the knowledge how many bits can be send over a link.

6. Final allocation of resource blocks, power and bits
   The actual allocation of resource blocks, power and bits is finalized by each AP on its own when the sizes of the subframes are known and the RSs know the initial data rate values of the BS-to-RS links.

The second concept is called distributed concept for reuse medium access since the concept is applicable if the reuse medium access scheme is considered and reuse groups have a cardinality larger than one. As motivated in the previous part of this section, the collection of the CQI values is based on a pilot phase, feed back phase and the actual transmission in the frame. Two assumptions are linked with the collection of the CQI values:

- Each AP transmits in a subframe of predefined subframe size in order to guarantee the same interfering stations during the whole pilot phase and whole actual transmission.

- The same power is allocated to each resource block in order to guarantee the same interference power in pilot phase and actual transmission.

The structure of the distributed concept for reuse medium access is changed compared to the distributed concept for orthogonal medium access. The structure illustrated in the flow chart given by Fig. 3.2 is as follows:

1. Initialization
   The initialization is the same as in the distributed concept for orthogonal medium access except that an additional assumption about the allocation of power is made. Since it is open in the beginning to which link the resource blocks are allocated and which interference and noise power occur, the power is allocated uniformly to all resource blocks.

2. Allocation of slots
   The allocation of slots is solved by the BS in the beginning since its solution is determined once and kept for all frames. The position of the RSs is assumed to be known. Since the UEs are nomadic or mobile, it is assumed that only the probability density function $\rho(x, y)$ is given, but not the actual positions.

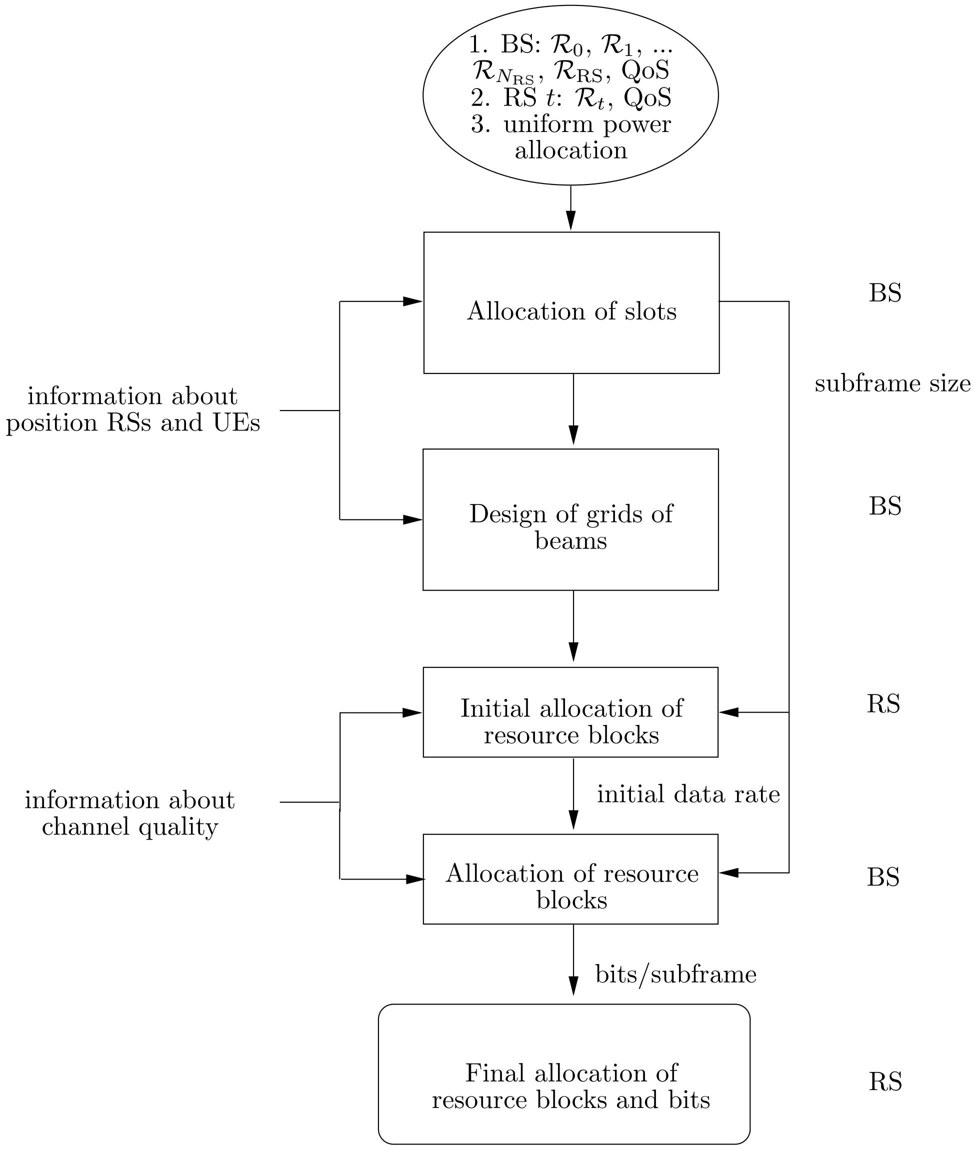

Figure 3.2. Flow chart of the distributed concept for reuse medium access.

The information about the position of the RSs and UEs is used to allocate more slots to APs which are expected to serve a large number of UEs. When this subproblem is solved, each AP knows the start and end of the subframes. This has an impact of the allocation of the resource blocks since a resource block has a fixed size and not an adaptive size as in the distributed concept for orthogonal medium access.

3. Design of grids of beams
   In addition to the distributed concept for orthogonal medium access, the design of grids of beams includes the problem to mitigate co-channel interference. Hence, a solution is required mitigating inter-beam and co-channel interference. When

this subproblem is solved, each AP knows which beamforming vectors are used for the time-frequency units.

4. Initial allocation of resource blocks by the RS
   Two differences occur compared to the distributed concept for orthogonal medium access. The first difference is that the allocation of power is predefined. Hence, only the initial allocation of resource blocks and not the initial allocation of resource blocks, power and bits must be solved. The allocation of power and bits simplifies to the selection of a modulation and coding scheme for the allocated resource blocks. The second difference is that the numbers of slots in the subframes are known. Hence, the formulation of the subproblem of initial allocation of resource blocks takes into account the number of slots. In the end of this step, the RS knows the number of bits which could be transmitted in a subframe and not only per slot. Hence, the initial data rate is defined as the number of bits per subframe in this concept. The initial data rate of each RS-to-UE link is sent to the BS.

5. Allocation of resource blocks by the BS
   As in the previous step, the allocation of power is predefined and the sizes of the subframes are known. Taking into account the initial data rate values received from all RSs in the cell, the BS solves the allocation of resource blocks. When the subproblem of the allocation of resource blocks is solved, the BS knows the data rate for each BS-to-RS and BS-to-UE link and the allocation is already finalized for these links. Due to the knowledge of the initial data rate values for the RS-to-UE links, it is impossible that an RS receives more bits per frame than it can forward. The RSs are informed about the bits transmitted in the frame over the BS-to-RS links.

6. Final allocation of resource blocks and bits
   The initial allocations of the RSs are finalized, i.e., it is checked wether a RS forwards as much bits as it receives from the BS. If necessary, the initial allocation of resource blocks is changed such that the data rates of the RS-to-UE links are matched to the BS-to-RS link.

3.4 Summary

Two resource allocation problems are defined in this Chapter. The definitions are related to the objective maximization of the minimum user rate and of the sum rate. The definition is valid for orthogonal and reuse medium access. It is motivated that optimum solutions of the problems are impractical because the computational complexity must be limited, the amount of signalling must be limited and CSI is hard to obtain. The distributed concept for orthogonal medium access and distributed concept for reuse medium access are introduced. The concepts differ since the former one is designed for the orthogonal medium access and the latter one for the reuse medium access. Each concept is applicable to both objectives. The distributed concept for orthogonal medium access decomposes a problem in the design of grids of beams, the initial allocation of resource blocks, power and bits solved by the RSs, the initial allocation of resource blocks, power and bits solved by the BS, the allocation of slots and the final allocation of resource blocks, power and bits. The distributed concept for reuse medium access decomposes a problem in the allocation of slots, the design of grids of beams, the initial allocation of resource blocks solved by the RSs, the allocation of resource blocks solved by the BS and the final allocation of resource blocks and bits.

Chapter 4
Distributed Concept for Orthogonal Medium Access

4.1 Introduction

In this chapter, the subproblems which must be solved in the distributed concept for orthogonal medium access are formulated and solutions are proposed. Each subproblem is treated according to the same structure:

- The open questions and the preliminaries related to the treated subproblem are formulated. The preliminaries include a list of parameters required to solve the subproblem.
- A non-adaptive algorithm is proposed. Such an algorithm is called non-adaptive since a solution is found without an adaptation to the listed parameters. Its computational complexity is negligible. The performance of a non-adaptive algorithm is rather limited. Its performance serves as a lower bound for the solution of the subproblem.
- An adaptive algorithm is proposed, where a solution adapted to the listed parameters is found. A mathematical description is given for the subproblem in order to motivate the derivation of the algorithms. The subproblem is formulated as a linear or non-linear integer program. Actually a wide range of algorithms exists finding the optimum solution, like the exhaustive search algorithm, dynamic programming algorithm or branch-and-bound algorithm. An overview is given in [NW99]. These algorithms have a huge computational complexity if the number of input parameters is large. However, the subproblem must be solved typically in milliseconds due to the limited coherence time of the channel. This restricts the acceptable computational complexity dramatically. Hence, novel algorithms are presented in this chapter. All algorithms are designed with respect to a low computational complexity, but some of them pay the low complexity with finding a suboptimum and not the optimum solution of the corresponding integer program.

Related to the structure mentioned above, the subproblem of the allocation of slots is an exception. The open questions, preliminaries and adaptive algorithms are given for the allocation of slots in this chapter, but a non-adaptive algorithm is not introduced. A non-adaptive algorithm leads to an fixed subframe size. This is treated in the distributed concept for reuse medium access, but is not treated in the distributed concept of orthogonal medium access.

The organization of the chapter is according to the successive order of the subproblems defined by the distributed concept for orthogonal medium access. Hence, Fig. 3.1 acts as a roadmap of this chapter. In Sections 4.2 to 4.6, the subproblems of the design of grids of beams, the initial allocation of resource blocks, power and bits by the RS and by the BS, the allocation of slots and the final allocation of resource blocks, power and bits are addressed successively. Section 4.7 summarizes the main contributions of this chapter.

4.2 Design of Grids of Beams

4.2.1 Open Questions and Preliminaries

In this section, the open questions and preliminaries are given for the non-adaptive and adaptive algorithms solving the design of grids of beams. According to Fig. 3.1, the design of grids of beams is solved by the BS. The open questions must be answered which beams are grouped in grids of beams and which grids of beams are allocated to time-frequency units. Since orthogonal medium access is assumed, the design of grids of beams is provided by the BS for each AP independently. For the non-adaptive and the adaptive algorithm, the BS knows the beamforming vectors of all APs in the cell. In combination with the antenna geometry, the BS is aware of the antenna pattern generated by a beamforming vector and representing the antenna gain depending on the direction of departure. A pattern with a clearly shaped main lobe is assumed, i.e., the magnitude of the main lobe is clearly larger than the one of the side lobes. Only the azimuth is considered.

For the adaptive algorithm, the BS requires an additional knowledge compared to the non-adaptive algorithm. This knowledge allows to use beams more often if receiving stations are expected in a particular direction although the exact positions of the UEs and the instantaneous CSI are not known. This knowledge is as follows:

- The BS knows the set $\mathcal{R}_{\mathrm{RS}}$ of RSs, the position of itself and of the RSs in $\mathcal{R}_{\mathrm{RS}}$. Hence, the directions from the BS to all RSs of the cell are known.
- The BS knows instantaneously which UEs are served by an AP t represented by the set $\mathcal{R}_t$. This knowledge is available at the BS and no additional effort must be spent to get this knowledge since the BS must address data of UEs having established two-hop connections to the correct RS. This knowledge is exploited to locate RSs which have a high traffic load.

- The BS knows the PDF $\rho(x, y)$ representing the distribution of the UEs in the cell. For instance, the PDF $\rho(x, y)$ can be measured during the operation of the relay network or in field trials in order to locate hot spot regions in which a large amount of UEs are connected to the relay network. The amount of data which is expected to be sent to these regions is estimated from the distribution of the UEs. For the sake of simplicity, it is assumed in this thesis that this amount is proportional to the number of UEs.
- The BS knows an appropriate pathloss model [Par00] well suited for the geographic environment in which the relay network operates. If the environment of the BS-to-RS links is different to the BS-to-UE and RS-to-UE links, the attenuation on the links is modelled by different pathloss models.

4.2.2 Non-Adaptive Design of Grids of Beams

In this section, the non-adaptive algorithm for the design of grids of beams is presented. The algorithm has been proposed in [GRB06b]. This proposal is summarized here briefly.

The algorithm is explained using exemplary beamforming vectors. According to Section 4.2.1, the pattern of each AP is known to the BS. A Uniform Circular Array (UCA) of omnidirectional elements is assumed for the APs. A coverage of 360° is assumed for the AP t. The coverage is ensured by $B_t = 12$ beamforming vectors. The main lobe direction of the first beamforming vector is 180°. The beamforming vector is found from the Dolph-Chebyshev approach [LL00], in which the beamforming vector is determined such that the width of the main lobe is minimized subject to a pre-defined side lobe attenuation. The pattern generated by the beamforming vector is illustrated in Fig. 4.1. The other 11 beamforming vectors are designed such that the main lobe direction of the illustrated beam is shifted from 180° to $0°, 30°, ...$ and 150° and from 180° to $210°, 240°, ...$ and 330°.

According to [GRB06b], B_t/G_t grids of beams are designed. Each beam is applied equally often. The B_t beams are grouped to grids of beams in a comb-like manner such that the inter-beam interference is minimized. The grids of beams are allocated to the time-frequency units in a periodical order. An example is given in Table 4.1 for $B_t = 12$ beamforming vectors and a size of $G_t = 2$ for the grids of beams. The beamforming vectors are identified by the main lobe direction of their beams. An extension of this strategy to $G_t = 1, 3, 4$ and 6 is straight forward. An extension to more than 6 beams is not beneficial since the patterns of neighbored beams overlab such that strong co-channel interference occurs. So far, the algorithm is only introduced for the exemplary beamforming vectors. However, the algorithm is applicable if beamforming vectors leading to clearly shaped main lobes are chosen.

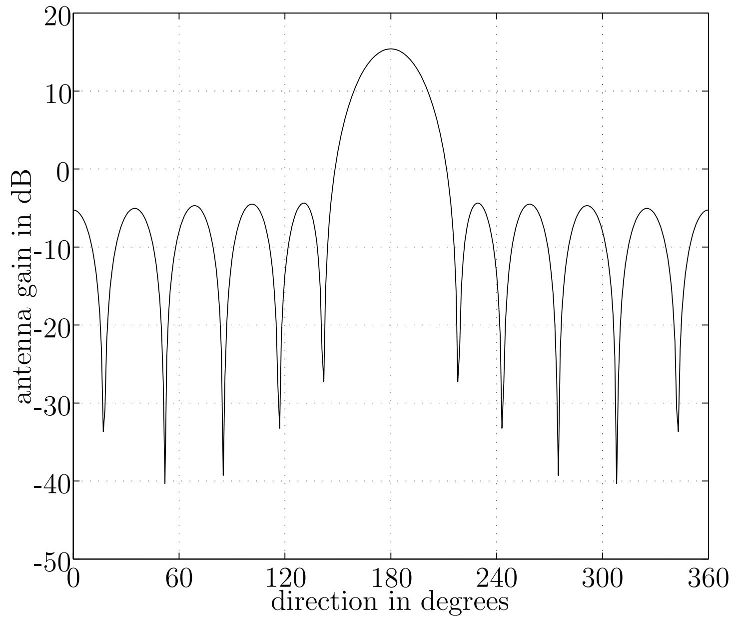

Figure 4.1. Pattern designed by Dolph-Chebyshev approach with main lobe direction in 180°.

Table 4.1. Example: beams used on time-frequency unit f for non-adaptive design of grids of beams for $B_t = 12$ and $G_t = 2$.

time-frequency unit f	1	2	3	4	5	6	7	...
first beam in grid of beams	0°	30°	60°	180°	210°	240°	0°	...
second beam in grid of beams	90°	120°	150°	270°	300°	330°	90°	...

4.2.3 Adaptive Design of Grids of Beams

In this section, the adaptive algorithm is introduced for the design of grids of beams. The algorithm adapts the design of grids of beams to the direction where UEs are expected. The two questions which beams are grouped in grids of beams and which grids of beams are allocated to the time-frequency units are answered in two steps. In the first step, the quantity representing how often a beam is applied among all time-frequency units is determined. This quantity is called demand value. In the second step, the grids of beams are determined. Taking into account the demand values, beams are grouped in grids of beams such that the interference is minimzed on average. The grids of beams are allocated to the time-frequency units. The adaptive design is applied to each AP independently since inter-beam interference, but not co-channel interference occurs. The adaptive design is applicable to the two objectives defined by (P1) and (P2) as the maximization of the minimum user rate and as the maximization of the sum rate, respectively.

The first step aims at maximizing the received signal power. The step is driven by the idea that those beams which are expected to be preferred by the UEs are used more often. This idea is proposed in [GRB06b] for the conventional network and extended to the relay network in the sequel.

How often a beam b of AP t is applied during a slot is given by the demand $D_{t,b}$. The demand values are determined under the constraint that

$$\sum_{b=1}^{B_t} D_{t,b} = G_t \cdot F \tag{4.1}$$

holds since all F time-frequency units shall be reused G_t times in space. The demand $D_{t,b}$ is determined as follows:

- A coverage area A_t is assigned to AP t. The coverage area of RS t is the area in which the UE r is using a two-hop connection most likely via the RS t and not most likely via another RS or a direct connection to the BS. The coverage area of the BS is the area in which the UE r is using most likely a direct connection to the BS. Based on an appropriate path loss model [Par00], the coverage area A_t can be determined depending on the path loss and the applied routing procedures.

- Each coverage area A_t is divided into B_t sectors, where each sector is assigned to a beam. The sector $A^{(t,b)}$ corresponds to the beam b of AP t. The sector $A^{(t,b)}$ is defined by the part of the coverage area in which the magnitude of the receive signal determined according to the pattern generated by beamforming vector $\mathbf{m}_{t,b}$ is expected to be larger than the magnitude of the other beams. Assuming that AP t covers 360°, a sector is limited by its two neighboring sectors and the border of the coverage area A_t.

- If the AP is an RS, the demand of beam b is determined by the probability that a UE is in sector $A^{(t,b)}$ and is given by

$$D_{t,b} = G_t \cdot F \cdot \frac{\int_{A^{(t,b)}} \rho(x,y)\ dxdy}{\int_{A_t} \rho(x,y) dxdy}, \tag{4.2}$$

where x and y are Cartesian coordinates describing the area of the cell. The denominator is required to normalize the distribution of the UEs in the sector to the coverage area of the RS t. The ratio of the integrals is positive and smaller than one. The factor $G_t F$ is introduced in order to fulfil (4.1).

- If the AP is the BS $t = 0$, the beam generated by beamforming vector $\mathbf{m}_{0,b}$ serves the UEs in $A^{(0,b)}$ and possibly one ore several RSs, if the main lobe of the beam b is directed to these RSs. These RSs are represented by the set $\mathcal{R}_{\mathrm{RS},b}$. Two summands are used to determine the demand of the BS. The first one corresponds to the number of UEs served by the RSs $\mathcal{R}_{\mathrm{RS},b}$ and given by $\sum\limits_{r\in\mathcal{R}_{\mathrm{RS},b}} |\mathcal{R}_r|$. The second one corresponds to the number of UEs with a direct connection and expected to be located in the sector $A^{(0,b)}$. The two summands are combined by normalizing the summands to the number N_{UE} of UEs in the cell in order to weight all UEs of the cell equally and in order to fulfill (4.1). The demand of beam b is given by

$$D_{t,b} = G_0 \cdot F \cdot \left(\frac{\sum\limits_{r\in\mathcal{R}_{\mathrm{RS},b}} |\mathcal{R}_r|}{N_{\mathrm{UE}}} + \frac{|\mathcal{R}_0|}{N_{\mathrm{UE}}} \frac{\int_{A^{(0,b)}} \rho(x,y)\ dxdy}{\int_{A_0} \rho(x,y)\ dxdy} \right). \tag{4.3}$$

  Note that the first summand is zero if the main lobe of the beam b of the BS is not directed to an RS. The term inside of the brackets is positive and smaller than one.

Since the demand values must be an integer value according to (4.1), a demand value is rounded to the next lower or larger integer value if it is not an integer number. The rounding has only a small impact if the number of time-frequency units F and the number G_t of beams applied in the grids of beams are large.

When the demand $D_{t,b}$ has been determined for each beam b, the second step in which the grids of beams are determined is processed. The second step aims at minimizing the interference while the demand values must be fulfilled. This step is explained by three parts: firstly a metric representing the interference is defined, secondly the minimization of the interference is formulated as an integer program and finally, the solution is formulated.

In order to describe whether a pair of beams generates a weak or strong interference if the pair is used on the same time-frequency unit, a metric is defined. Since the actual interference is not known, a simplified model of the occurring interference is used and the metric is defined as an average interference. Since orthogonal medium access is considered, only inter-beam interference appears while co-channel interference does not occur. Firstly, the interference at a single position is modelled, then the value averaged over all positions in a sector is introduced. The interference at the position (x, y) in the sector $A^{(t,b)}$ is modelled by the product of three terms. The first one is the transmit power $P^{\mathrm{TX}}_{(t,b')}$ if beam b' is applied, where $b' = 1, 2, ...B_t$. For simplicity, it is assumed that the power of the AP t is uniformly allocated to the resource blocks. The actual power value is determined by the APs after the design of grids of beams according to the illustration of the concept in Fig. 3.1. The second term is the antenna gain determined according to the pattern and denoted by $G_{(t,b')}(x, y)$. The last one corresponds to the attenuation of the channel from AP t to (x, y) and is denoted by $PL_{(t)}(x, y)$. The attenuation can be found by the pathloss model chosen according to the geographic environment in which the relay network operates. The interference $I_{(t,b')}(x, y)$ generated by beam b' of AP t is modelled by

$$I_{(t,b')}(x, y) = PL_{(t)}(x, y) \cdot G_{(t,b')}(x, y) \cdot P^{\mathrm{TX}}_{(t,b')}, \tag{4.4}$$

where (x, y) is in $A^{(t,b)}$. Equation (4.4) is used to define an average interference caused by beam b' in sector $A^{(t,b)}$. If the AP t is an RS, the average interference is

$$I^{(t,b)}_{(t,b')} = \int\int_{A^{(t,b)}} I_{(t,b')}(x, y)\rho(x, y)\ dxdy, \tag{4.5}$$

where the average value is found by weighting the interference at position (x, y) with the PDF representing the probability that a UE of the cell is at the position (x, y). In practice, the interference can be approximated by considering a finite number of positions within the sector $A^{(t,b)}$.

If the AP t is the BS, the definition of the average interference must consider that the receiving stations of the BS are UEs and RSs. The average interference $I^{(0,b)}_{(0,b')}$ is generated by beam b' of the BS. The average interference occurs at the UEs in the coverage area $A^{(0,b)}$ and at the RSs $\mathcal{R}_{\mathrm{RS},b}$ if the main lobe of beam b is directed to these RSs. The interference related to the UEs is obtain by the integral over the PDF representing the distribution of the UEs. The interference related to the RSs is obtained by averaging the interference of the RSs. The interference related to a single RS is determined based on the known position of the RS. The RS r, where $r \in \mathcal{R}_{\mathrm{RS},b}$, is assumed to be placed at (x_r, y_r). The interference is $I_{(0,b')}(x_r, y_r)$ for (x_r, y_r) according to (4.4). The average interference $I^{(0,b)}_{(0,b')}$ is defined by

$$I^{(0,b)}_{(0,b')} = \frac{|\mathcal{R}_0|}{N_{\mathrm{UE}}} \int\int_{A^{(0,b)}} I_{(0,b')}(x, y)\rho(x, y)\ dxdy + \sum_{r \in \mathcal{R}_{\mathrm{RS},b}} \frac{|\mathcal{R}_r|}{N_{\mathrm{UE}}} I_{(0,b')}(x_r, y_r). \tag{4.6}$$

The two summands corresponding to the interference received by the UEs and the RSs are linearly combined. The weights are chosen in order to obtain an average value in which all UEs of the cell are weighted equally.

In total, B_t^2 values of average interferences are defined by (4.5) or (4.6) for AP t. These values are written in an interference matrix. The interference matrix is called $\acute{\mathbf{A}} \in \mathbb{R}^{B_t \times B_t}$. The matrix represents the metric required for the design of grids of beams for AP t.

Based on the metric, the problem of grouping of beams in grids of beams and assigning them to the time-frequency units is formulated in mathematical terms as an integer program. The indicator variable $v_{t,b,k,n}$ defined in (2.2) is simplified to $v_{t,b,k}$. The index n is omitted since only subframe n is considered in the following. The variables $v_{t,b,k}$ for AP t, for all b and for all k are written in the vector $\mathbf{v} \in \{0,1\}^{B_t F \times 1}$. Since the matrix $\acute{\mathbf{A}}$ describes only the interference if beams are applied for the same time-frequency unit, an extension of the description is required for all time-frequency units. This extension is provided by the block diagonal matrix $\mathbf{A} \in \mathbb{R}^{B_t F \times B_t F}$. The blocks of the diagonal have a size of $B_t \times B_t$ and are filled with the matrix $\acute{\mathbf{A}}$. Outside of the diagonal, the elements of the matrix are zero. The objective function of the integer program is chosen such that the interference is mitigated for each time-frequency unit and is written as minimizing the sum of all average interference values. The minimization is subject to the constraint (2.3) which claims that G_t beams are allocated to each time-frequency unit and to the constraint that the demand value $D_{t,b}$ must be fulfilled for each beam. The integer program is given by

$$\min_{\mathbf{v}} \mathbf{v}^{\mathrm{T}} \mathbf{A} \mathbf{v} \tag{4.7a}$$

subject to:

$$\sum_{b=1}^{B_t} v_{t,b,f} = G_t, \qquad 1 \leq f \leq F, \tag{4.7b}$$

$$\sum_{f=1}^{F} v_{t,b,f} = D_{t,b}, \qquad 1 \leq b \leq B_t, \tag{4.7c}$$

$$v_{t,b,f} \in \{0,1\}. \tag{4.7d}$$

A solution of the integer program is required in real-time applications in the relay network. Finding the optimum solution of the integer program is too complex in real-time applications. Since the product of the number of time-frequency units F and the number B_t of available beams is rather large, the number of assignments variable given in the vector $\mathbf{v} \in \{0,1\}^{B_t F \times 1}$ becomes also large. As stated in Section 4.1, large integer programs are hard to solve in real time applications. Hence, a suboptimum algorithm is proposed, where beams are allocated sequentially. The algorithm is shown in Fig. 4.2 and consists of two steps. In the first step, the set $\mathcal{G}_f$ is initialized for each time-frequency unit f. The set $\mathcal{G}_f$ represents all beams assigned to f, i.e. a grid of beams. In the second step, beams are allocated to time-frequency units sequentially. The beam

Step 1: *Initialization*

$\mathcal{G}_f := \{\}$ for all f

Step 2: *Design grids of beams*

for $f = 1$ to F **do**

for $g = 1$ to G_t **do**

$b^* := \underset{b}{\operatorname{argmax}}\{D_{b,t}\}$

Choose best fitting time-frequency unit according to (4.8)

$\mathcal{G}_{f^*} := \mathcal{G}_{f^*} \cup \{f^*\}$

$D_{(t,b^*)} := D_{(t,b^*)} - 1$

Figure 4.2. Adaptive design of grids of beams.

b^* which has the maximum demand value is chosen. The beam b^* is allocated to the time-frequency unit f^*. The time-frequency unit f^* is found by searching the time-frequency unit such that the beam b^* causes the lowest interference in the sector of the already allocated beams and such that the lowest interference occurs in the sector $A^{(t,b^*)}$. The time-frequency unit f^* is searched among all units to which not yet G_t beams are assigned. These time-frequency units are represented by the set $\mathcal{F}$. The time-frequency unit f^* is found by

$$f^* = \underset{\mathcal{F}}{\operatorname{argmin}} \left(\sum_{g \in \mathcal{G}_f} I_{(t,b^*)}^{(t,g)} + I_{(t,g)}^{(t,b^*)} \right). \tag{4.8}$$

The beam b^* is assigned to $\mathcal{G}_{f^*}$ and its demand value D_{t,b^*} is decreased by one. The sequential allocation is repeated FG_t times such that a grid of beams is allocated to each time-frequency unit.

The computational complexity required for the adaptive algorithm for the design of grids of beams is given in the following. In this thesis, the computational complexity is represented by three terms. The first term is the number of multiplications and divisions. Both operations are included in the same term since divisions have the same complexity as multiplications if they are implemented efficiently, e.g., using Newton's method [BV04]. The second term is the number of additions and subtractions. Here both operations are included in the same term since subtractions are considered as additions taking into account the algebraic sign. The third term is the number of searches for a maximum or minimum. In addition to the number of searches for a maximum or minimum, the size of the searched set is given since the computational complexity of the search depends on the size of the given set. For the adaptive design of grids of beams, it is assumed that the PDF $\rho(x, y)$, the number F of time-frequency units, the number G_t of applied beams and the number B_t of beamforming vectors are invariant over the time for all APs, only the number of UEs and their assignment is time-variant. Then, the computational load is caused by the calculation of the demand values of the BS according to (4.3), the calculation of the average interference values of the BS according to (4.6) and the algorithm depicted in Fig. 4.2. The calculation of the demand values of the RSs according to (4.2) and of the average interference values of the RSs according to (4.5) are not time-variant. The computational complexity is given in Table 4.2 for solving the subproblem of the design of grids of beams for all APs in the cell. Since the search of time-frequency unit f^* in (4.8) depends on the instantaneous allocation, an upper bound is given for the computational complexity of (4.8). It is assumed that the time-frequency unit f^* is always searched among all F time-frequency units and that $G_t - 1$ beams are allready allocated in the set $\mathcal{G}_f$.

Table 4.2. Computational complexity of the adaptive algorithm for the design of grids of beams.

Multiplications/divisions	$3B_0 + B_0 \sum_{b=1}^{B_0} (\lvert\mathcal{R}_{\mathrm{RS},b}\rvert + 2)$
Additions/subtractions	$\sum_{\text{all } t} (FG_t + (2G_t - 3)F^2 G_t) + (B_0 + 1) \sum_{b=1}^{B_0} \lvert\mathcal{R}_{\mathrm{RS},b}\rvert$
Search maximum/minimum	for each t: search $G_t F$ times in set of size B_t for each t: search $G_t F$ times in set of size F

4.3 Initial Allocation of Resource Blocks, Power and Bits by the RS

4.3.1 Open Questions and Preliminaries

In this section, the open questions and preliminaries are given for the initial allocation of resource blocks, power and bits, where the initial allocation is solved by each RS individually. For the RS t, the questions must be answered which resource blocks are allocated to the links (t, r), where $r \in \mathcal{R}_t$, and which power is allocated to a resource block in order to apply a certain modulation and coding scheme. The RS t knows:

- the CQI values defined as the instantaneous channel gain values and noise power values for each time-frequency unit and for each link (t, r), where $r \in \mathcal{R}_t$.

- the SINR threshold $\gamma_{\epsilon,r}$ for each number ϵ of bits, where $\epsilon \in \mathcal{E}$, and for each receiving UE $r \in \mathcal{R}_t$ in order to guarantee the bit error probability $BEP_{t,r}$ for each link (t, r).

- the minimum data rate $R_{\min,t,r}$ for each link (t, r) if the objective is to maximize the sum rate as claimed in problem (P2).

The subproblem must be solved without knowing the subframe sizes in a frame and the data rate transmitted from the BS to the RS t.

4.3.2 Non-Adaptive Allocation of Resource Blocks, Power and Bits by the RS

In this section, the non-adaptive algorithms allocating resource blocks, power and bits are presented. Based on a uniform power allocation, resource blocks are allocated to the links by a weighted round robin algorithm which considers the bits which can be transmitted in the already allocated resource blocks. At first, the allocation of power is treated, then the selection of the modulation and coding scheme and finally, the allocation of the resource blocks.

Firstly the power is allocated. Since it is not known initially to which link (t, r) the resource block k is allocated, each resource block is treated equally. Hence, the total transmit power P_t is allocated to $K_t = G_t F$ resource blocks uniformly and is independent of the receiving station. The power allocated to resource block k is

$$p_{t,k} = P_t / K_t. \tag{4.9}$$

Using the CQI values, the SINR value $\gamma_{r,k}$ is determined for each link (t, r), where $r \in \mathcal{R}_t$ and for each resource block k. The SINR values are used to choose the modulation and coding schemes for all links and for all resource blocks. The number $\epsilon_{r,k}$ of bits allocated to a slot of resource block k and link (t, r) is given by

$$\epsilon_{r,k} = \begin{cases} \epsilon_1 & \text{for } 0 \leq \gamma_{r,k} < \gamma_{\epsilon_1,r} \\ \epsilon_2 & \text{for } \gamma_{\epsilon_1,r} \leq \gamma_{r,k} < \gamma_{\epsilon_2,r} \\ \vdots & \\ \epsilon_{|\mathcal{E}|} & \text{for } \gamma_{\epsilon_{|\mathcal{E}|-1},r} \leq \gamma_{r,k} < \infty, \end{cases} \tag{4.10}$$

where ϵ_1, ϵ_2, ...$\epsilon_{|\mathcal{E}|}$ denote the elements of the set $\mathcal{E}$. The uniform power allocation and the presented bit allocation is called non-adaptive allocation of power and bits.

The non-adaptive algorithm for the allocation of resource blocks is the weighted round robin algorithm [SBS99] applied by the RS t. Resource blocks are allocated successively to the links. The resource block k is allocated to the link (t, r) which has the minimum value of the product of the number of already allocated resource blocks and the weight w_r. The weights w_r are defined for each $r \in \mathcal{R}_t$. If the objective of the concept is given by (P1) as maximizing the minimum user rate, the weight w_r is

$$w_r = 1 \tag{4.11}$$

in order to treat all UEs equally. If the objective of the concept is (P2), the weights shall ensure that the UEs demanding a large minimum data rate are allocated a large number of resource blocks. For this case, the weight w_r is

$$w_r = \frac{R_{\min,t,r}}{\sum\limits_{r \in \mathcal{R}_t} R_{\min,t,r}}. \tag{4.12}$$

When the non-adaptive, initial allocation is finished, the RS knows the maximum number of bits which could be transmitted to the UE r in one slot. This value is called initial data rate value $\hat{R}_{t,r}$. The initial data rate value $\hat{R}_{t,r}$ is sent by the RS to the BS for all $r \in \mathcal{R}_t$ in order to support the BS in its allocation.

4.3.3 Adaptive Allocation of Resource Blocks, Power and Bits by the RS Aiming at Maximizing the Minimum User Rate

In this section, the subproblem of the allocation of resource blocks, power and bits is treated for RS t if the considered objective of the concept is defined by (P1) as the maximization of the minimum user rate. To reduce the complexity of the subproblem, the subproblem is divided into two smaller ones. At first, the allocation of resource blocks is solved in order to provide each link suitable resource blocks. Assuming a uniform power allocation, bits are mapped to each resource blocks and a first initial data rate value is determined for each link. Then, the allocation of power and bits is solved adaptively in order to increase the initial data rate values. Both subproblems are treated in the following according to the same structure: at first, a subproblem is formulated as an integer program, then an adaptive algorithm solving the subproblem and requiring a low computational complexity is proposed.

Firstly, the allocation of resource blocks is treated. Since the power is not allocated to resource blocks yet, the assumption is made that the power is allocated uniformly among the resource blocks. Taking into account the CQI values defined as channel gain and noise power values, the RS t is aware of the SINR value $\gamma_{r,k}$ for each of the K_t resource blocks and for each $r \in \mathcal{R}_t$. By means of (4.10), a number $\epsilon_{r,k}$ of bits is assigned to each SINR value $\gamma_{r,k}$. These bits will be transmitted in a slot of the resource block k if k is allocated to UE r. In order to describe the subproblem, the assignment variable $u_{t,r,\epsilon,k,n}$ defined in (2.6) is simplified to $u_{r,k}$ since only one RS t and one subframe n is considered and $\epsilon_{r,k}$ bits are assumed for UE r and resource block k. The objective is to maximize the lowest data rate $R_{t,r}$ defined in (2.22) for the link $r \in \mathcal{R}_t$ and given by the minimum of the data rates searched over all $r \in \mathcal{R}_t$. The data rate $R_{t,r}$ is weighted by w_r. The weight w_r is used in order to give the most general formulation of the subproblem and simplifies the formulations in the following sections. The usage of the weights allows that single UEs are preferred. If the RS solves the allocation of resource blocks with the aim related to (P1) adaptively, each UE is treated equally and the weight w_r is defined as

$$w_r = 1 \tag{4.13}$$

for all $r \in \mathcal{R}_t$. The allocation of resource blocks is subject to the constraints defined by (2.10) and (2.11). The constrains describe that at most one resource block of the same time-frequency unit is allocated to a link and that each resource block is allocated exclusively to one link, respectively. The subproblem of the allocation of resource blocks is given by the following integer program:

$$\max_{u_{r,k}} \min_{r \in \mathcal{R}_t} w_r \sum_{k=1}^{K_t} u_{r,k} \epsilon_{r,k} \tag{4.14a}$$

subject to:

$$\sum_{k=(f-1)G_t+1}^{fG_t} u_{r,k} \leq 1, \qquad 1 \leq f \leq F,\ r \in \mathcal{R}_t, \tag{4.14b}$$

$$\sum_{r \in \mathcal{R}_t} u_{r,k} = 1, \qquad 1 < k \leq K_t, \tag{4.14c}$$

$$u_{r,k} \in \{0,1\}, \qquad 1 < k \leq K_t,\ r \in \mathcal{R}_t. \tag{4.14d}$$

Finding the optimum solution of this integer program is rather complex. Without considering spatial multiplexing and neglecting (4.14b), this integer program is formulated in [RC00]. Since the optimum solution of the integer program requires a large computational effort, a greedy algorithm is proposed in [RC00]. In this thesis, the algorithm proposed in [RC00] is extended and matched to the considered application. A pseudo code of the algorithm is illustrated in Fig. 4.3. The idea is that resource blocks are allocated successively. In each iteration, a resource block is allocated to the link identified by the lowest number of bits carried by the resource blocks allocated in

Step 1: *Initialization*

$\hat{R}_{t,r} := 0$ for all $r \in \mathcal{R}_t$

$\mathcal{K} := \{1, 2, ..., K_t\}$

$\mathcal{K}_r := \{\}$ for all $r \in \mathcal{R}_t$

Step 2: *Allocate first resource block to each link*

for all $r \in \mathcal{R}_t$ **do**

$$k^* := \underset{k \in \mathcal{K}}{\operatorname{argmax}}\{\gamma_{r,k}\}$$

$$\mathcal{K}_r := \mathcal{K}_r \cup \{k^*\}$$

$$\hat{R}_{t,r} := \hat{R}_{t,r} + \epsilon_{r,k^*}$$

$$k^* \mapsto f^*$$

$$\gamma_{r,k} := -\infty \text{ for } (f^* - 1)G_t + 1 \leq k \leq f^* G_t$$

$$\mathcal{K} := \mathcal{K} \setminus \{k^*\}$$

Step 3: *Increase data rates*

while $\mathcal{K} \neq \{\}$ **do**

$$r^* := \underset{r \in \mathcal{R}_t}{\operatorname{argmin}}\{w_r \hat{R}_{t,r}\}$$

$$k^* := \underset{k \in \mathcal{K}}{\operatorname{argmax}}\{\gamma_{r^*,k}\}$$

$$\mathcal{K}_{r^*} := \mathcal{K}_{r^*} \cup \{k^*\}$$

$$\hat{R}_{t,r^*} := \hat{R}_{t,r^*} + \epsilon_{r^*,k^*}$$

$$k^* \mapsto f^*$$

$$\gamma_{r^*,k} := -\infty \text{ for } (f^* - 1)G_t + 1 \leq k \leq f^* G_t$$

$$\mathcal{K} := \mathcal{K} \setminus \{k^*\}$$

Figure 4.3. Resource block allocation algorithm aiming at maximizing the minimum user rate according to the objective of (P1).

previous iterations. The algorithm is structured in three steps. In the first step, the initial data rate value $\hat{R}_{t,r}$ is set to zero, the sets $\mathcal{K}$ and $\mathcal{K}_r$ are defined in order to describe the resource blocks available for the allocation and the resource blocks allocated to r, respectively. In the second step, a first resource block is allocated to each receiver successively. For receiver r, the resource block k^* with the maximum SINR value γ_{r,k^*} is found and assigned to $\mathcal{K}_r$. The initial data rate value $\hat{\mathcal{R}}_{t,r}$ is increased by ϵ_{r,k^*} bits. The time-frequency unit f^* which is used by the resource block k^* is determined. This operation is denoted by $k^* \mapsto f^*$. The SINR values of all resource blocks using f^* are set to a small value for r in order to prevent that more than one resource block of the same time-frequency unit f^* is allocated to link (t, r). Then, the resource block k^* is removed from set $\mathcal{K}$ in order to prevent that k^* is allocated twice. In the third step, the remaining resource blocks of set $\mathcal{K}$ are allocated by a loop. In each iteration, the receiver r^* with the minimum weighted initial data rate value $w_{r^*}\hat{R}_{t,r^*}$ is selected. A resource block k^* is allocated to the receiver r^* by processing the same operations as in the second step. The algorithm ends if all resource blocks are allocated.

The computational complexity required for the adaptive algorithm is given in Table 4.3. The computational complexity is a function of the number F of time-frequency units, the number G_t of beams and the number $|\mathcal{R}_t|$ of assigned UEs. The uniform allocation of power and bits assumed in the formulation of the subproblem of the allocation of resource blocks is not included since it is also required for the non-adaptive allocation of resource blocks, power and bits. The value of the multiplications is given for the worst case assumption that $w_r \neq 1$ for each receiving station r. If $w_r = 1$, multiplications by the factor w_r can be omitted.

Table 4.3. Computational complexity of the resource block allocation algorithm aiming at maximizing the minimum user rate.

Multiplications/divisions	$FG_t - 1$
Additions/subtractions	FG_t
Search maximum/minimum	search FG_t times in set of size FG_t, search $FG_t - |\mathcal{R}_t|$ times in set of size $|\mathcal{R}_t|$

For each link (t,r), where $r \in \mathcal{R}_t$, the initial data rate value $\hat{R}_{t,r}$ is found by the algorithm depicted in Fig. 4.3, where it is assumed that the power is allocated uniformly. Solving the allocation of power and bits adaptively promises higher initial data rate values. The formulation of the subproblem of the allocation of power and bits as an integer program and its adaptive solution is given in the following.

The corresponding subproblem is defined as follows. The assignment variable $u_{t,r,\epsilon,k,n}$ defined in (2.6) is simplified to $u_{k,\epsilon}$ since the resource block k is uniquely allocated to the link (t,r) and one subframe n is considered. If ϵ bits are transmitted in a slot of resource block k, the power $p_{t,k}$ must be allocated according to (3.3). Since inter-beam interference exists in the general case, the power allocated to resource block k based on time-frequency unit f is a function of the allocated power on all resource blocks k', where $(f-1)G_t \leq k' \leq fG_t$ and $k' \neq k$. Taking into account (3.3), the power $p_{t,k}$ is given by

$$p_{t,k} = \alpha_{t,r,k}^{-1}\gamma_\epsilon \left(Z_{r,k} + \sum_{\substack{k'=(f-1)G_t+1 \\ k' \neq k}}^{fG_t} \sum_{\epsilon \in \mathcal{E}} p_{t,k'} u_{k,\epsilon} \right), \tag{4.15}$$

where $\alpha_{t,r,k}$ and $Z_{r,k}$ denote the channel gain value and noise power value of link (t, r) and resource block k. In (4.15), the first sum corresponds to all interfering resource blocks, the second sum corresponds to the modulation and coding scheme of the set $\mathcal{E}$. For the definition of the objective function of the integer program, the data rate is weighted by w_r defined by (4.13) like in the previous subproblem. Taking into account the constraints given by (2.7) and (2.11) and describing that the sum of the transmitted power is limited and a modulation and coding scheme must be allocated to each resource block, the subproblem of the allocation of power and bits is defined as:

$$\max_{u_{k,\epsilon}} \min_{r\in\mathcal{R}_t} w_r \sum_{k\in\mathcal{K}_r} \sum_{\epsilon\in\mathcal{E}} u_{k,\epsilon}\epsilon \tag{4.16a}$$

subject to:

$$\sum_{k=1}^{K_t} \sum_{\epsilon\in\mathcal{E}} p_{t,k} u_{k,\epsilon} \le P_t, \tag{4.16b}$$

$$\sum_{\epsilon\in\mathcal{E}} u_{k,\epsilon} = 1, \qquad 1 \le k \le K_t, \tag{4.16c}$$

$$u_{k,\epsilon} \in \{0, 1\}, \qquad 1 < k \le K_t,\ \epsilon \in \mathcal{E}. \tag{4.16d}$$

According to (4.15) and (4.16b), the power $p_{t,k}$ is a quadratic and neither a convex nor a concave function depending on the assignment variables [NW99] if inter-beam interference exists, i.e., if $G_t > 1$. The optimum solution is only found with a high computational complexity. Here, a suboptimum algorithm is proposed with a low complexity and applicable to real time applications.

A bit and power loading algorithm is introduced in the following. Taking into account that the resource blocks are allocated to the links, power and bits allocated to the resource blocks are initially set to zero. If the modulation and coding scheme providing the next higher number of bits is applied for a resource block, an increment of power is required. The increment of power is determined for each resource block. The bits are allocated to the resource block which requires the minimum increment of power. Bits are allocated as long as the power limit is not exceeded. At first, the power required for a modulation and coding scheme is derived, then the algorithm is presented in detail.

In order to derive the increment of power, a deeper analysis of the allocated power given by (4.15) is necessary. Assume that ϵ_k bits are allocated to each resource block k and the SINR threshold γ_{ϵ_k} is met. The resource blocks using the same time-frequency unit f are coupled by the inter-beam interference. Since all of them fulfill their SINR thresholds, equation

$$\alpha_{t,r,k} p_{t,k} - \gamma_{\epsilon_k} \sum_{\substack{k'=(f-1)G_t+1 \\ k' \neq k}}^{fG_t} \alpha_{t,r,k'} p_{t,k'} = \gamma_{\epsilon_k} Z_k \tag{4.17}$$

holds for $k = (f-1)G_t+1, (f-1)G_t+2, ...fG_t$. These G_t equations are written in vector matrix notation. The channel gain values including their algebraic signs and the noise power values multiplied by the SINR thresholds are represented by $\mathbf{A}_f \in \mathbb{R}^{G_t \times G_t}$ and $\mathbf{Z}_f \in \mathbb{R}^{G_t \times 1}$, respectively. The transmit power values are merged in $\mathbf{p}_f(\epsilon_k) \in \mathbb{R}^{G_t \times 1}$. The variable ϵ_k indicates that the vector contains the power values under the constraint that ϵ_k bits are considered for resource block k. The equations defined by (4.17) are rewritten as

$$\mathbf{A}_f \mathbf{p}_f(\epsilon_k) = \mathbf{Z}_f \Leftrightarrow \mathbf{p}_f(\epsilon_k) = \mathbf{A}_f^{-1} \mathbf{Z}_f. \tag{4.18}$$

Since only positive transmit power values can be allocated, $\mathbf{p}_f(\epsilon_k)$ is only applicable if all elements of $\mathbf{p}_f(\epsilon_k)$ are larger than or equal to zero. The power values $\mathbf{p}_f(\epsilon_k + \Delta\epsilon)$ required to transmit $\epsilon_k + \Delta\epsilon$ bits are also calculated according to (4.18). The increment of power denoted by $\Delta\mathbf{p}_k$ is given by

$$\Delta\mathbf{p}_k = \mathbf{p}_f(\epsilon_k + \Delta\epsilon) - \mathbf{p}_f(\epsilon_k) \tag{4.19}$$

if all transmit power values are larger than or equal to zero. Otherwise, the allocation of $\Delta\epsilon$ bits is not possible as denoted by infinitely large values for the elements of $\Delta\mathbf{p}_k$. The 2-norm $||\Delta\mathbf{p}_k||$ represents the additional power which must be allocated if the number of bits loaded to resource block k is increased by $\Delta\epsilon$ bits.

Based on the derivation of the increment of power, the total algorithm as illustrated in Fig. 4.4 is presented. The algorithm consits of two steps. In the first step, the initial data rate value $\hat{R}_{t,r}$ and the number ϵ_k of allocated bits are set to zero for all r and k. The sum p of the allocated power values is also zero. The increment $||\Delta\mathbf{p}_k||$ of power is determined for each resource block k. In the second step, the sum p is increased as long as the limit P_t is not exceeded. At first, the receiver r^* with the minimum weighted initial data rate value $w_{r^*}\hat{R}_{t,r^*}$ is found. The resource block k^* is searched in the set $\mathcal{K}_{r^*}$ of allocated resource blocks. If the limit P_t allows, additional $\Delta\epsilon$ bits are allocated, the initial data rate $\hat{R}_{t,r^*}$ is increased and the increments are updated for all resource blocks of the same time-frequency unit denoted by f^*. The algorithm ends if the power not allocated so far is too small in order to increase the number of bits for any resource block.

The computational complexity required for the adaptive algorithm is given in Table 4.4. The computational complexity includes the calculation of $\Delta\mathbf{p}_k$ by solving the

Step 1: *Initialization*

$\hat{R}_{t,r} := 0$ for all $r \in \mathcal{R}_t$

$\epsilon_k := 0$ for all $k \in \mathcal{K}_r$ and for all $r \in \mathcal{R}_t$

$p := 0$

calculate $||\Delta\mathbf{p}_k||$ according to (4.19) for all k

Step 2: *Increase data rates*

while $p \leq P_t$ **do**

$$r^* := \underset{r \in \mathcal{R}_t}{\operatorname{argmin}}\{w_{r^*}\hat{R}_{t,r^*}\}$$

$$k^* := \underset{k \in \mathcal{K}_{r^*}}{\operatorname{argmin}}\{||\Delta\mathbf{p}_k||\}$$

$p := p + ||\Delta\mathbf{p}_{k^*}||$

if $p \leq P_t$, **do** $\epsilon_{k^*} := \epsilon_{k^*} + \Delta\epsilon$ and $\hat{R}_{t,r^*} := \hat{R}_{t,r^*} + \Delta$

$k^* \mapsto f^*$

update $||\Delta\mathbf{p}_k||$ according to (4.19) for $(f^* - 1)G_t + 1 \leq k \leq f^*G_t$

if $\epsilon_{k^*} = \max\{\mathcal{E}\}$, **do** $||\Delta\mathbf{p}_{k^*}|| := \infty$

Figure 4.4. Bit and power loading algorithm aiming at maximizing the minimum user rate according to the objective of (P1).

system of linear equation given by (4.19). It is assumed that a system of linear equations is solved by a matrix inversion, where an inversion requires $(G_t^3 + 3G_t^2)/2$ multiplications and $(G_t^3 + G_t^2)/2$ additions [Hun07]. Since methods like the Cholesky decomposition may be more efficient [Hun07], the presented values of the computational complexity serve as an upper bound. The actual computational complexity depends on the number of allocated bits. Here, the worst case assumption is made that the maximum number of bits is allocated to each resource block. In addition to the values given in Table 4.4, $(|\mathcal{E}| - 1)FG_t$ square roots operations are required in order to determine the 2-norm of $\Delta\mathbf{p}_k$.

Table 4.4. Computational complexity of the bit and power loading algorithm aiming at maximizing the minimum user rate.

Multiplications/divisions	$(\lvert\mathcal{E}\rvert - 1)FG_t(\frac{1}{2}G_t^3 + \frac{7}{2}G_t^2 + G_t + 1) - 1$
Additions/subtractions	$(\lvert\mathcal{E}\rvert - 1)FG_t(\frac{1}{2}G_t^3 + \frac{1}{2}G_t^2 + G_t + 2)$
Search maximum/minimum	search $(\lvert\mathcal{E}\rvert - 1)FG_t$ times in set of size $\lvert\mathcal{R}_t\rvert$, search $(\lvert\mathcal{E}\rvert - 1)FG_t$ times in set of size F

When the allocation of power and bits is solved, the RS knows the initial data rate value $\hat{R}_{t,r}$ for each $r \in \mathcal{R}_t$. All initial data rate values are sent by the RS to the BS in order to support the allocation provided by the BS.

4.3.4 Adaptive Allocation of Resource Blocks, Power and Bits by the RS Aiming at Maximizing the Sum Rate

In this section, the subproblem of the allocation of resource blocks, power and bits is treated for the RS t if the considered objective of the concept is to solve (P2) defined as the maximization of the sum rate while the minimum data rate $R_{\min,t,r}$ is guaranteed for each $r \in \mathcal{R}_t$. The objective is achieved by two steps.

1. The subproblem is solved aiming at offering the minimum data rate $R_{\min,t,r}$.

2. The subproblem is solved aiming at maximizing the sum rate.

The section is structured as follows. At first, the two steps are motivated. Then both steps are treated separately. In both steps, firstly the allocation of resource blocks is treated, then the allocation of power and bits. As motivated in Section 4.3.3, first the allocation of resource blocks is solved in order to provide links resource blocks with high channel qualities, then the allocation of power and bits is solved in order to improve the initial data rate values.

Achieving the objective given by (P2) in two steps is motivated as follows. The sizes of the subframes are determined after the initial allocation of the RS according to Fig. 3.1. Hence, the size of any subframe is not known to the RS t yet. If a resource block is allocated to a link, the total number of bits transmitted by the resource block is not predictable. Hence, the number of resource blocks required to be allocated to link (t, r) such that the minimum data rate $R_{\min,t,r}$ is achieved cannot be determined. The division into two steps has the advantage that one subframe is reserved for the RS t for each step. The subproblem of the allocation of resource blocks, power and bits is proposed to be solved independently for each subframe with different objectives. In the first subframe, the objective is to achieve the minimum data rates since the minimum data rates must be provided. In the second subframe, the strategy of maximizing the sum rate is pursued as well as possible. In both subframes, the subproblem is solved where only one slot is considered in both steps. Since only one slot is considered, the optimization is related to initial data rate values representing the bits per single slot. If further slots are allocated to the RS t, the data rate value of the link (t, r) is a linear function of the initial data rate values and the number of allocated slots. Thus, the decision whether the minimum data rate is achieved is made if the subproblem allocation of slots is solved.

For the first subframe allocated to RS t, firstly the allocation of resource blocks is shown, then the allocation of power and bits.

Initially, a uniform allocation of the power is assumed such that the number of bits $\epsilon_{r,k}$ is given for each k and r. The subproblem of the allocation of resource is given by (4.14) as shown in the following. The objective related to the subproblem is that the size of the first subframe is minimized while the minimum data rate is achieved in order to provide as many slots as possible for the second subframe which is used to maximize the sum rate. If all links request the same minimum data rate, the resource blocks, power and bits must be allocated such that all links achieve the same initial data rate. Otherwise, the link achieving the minimum initial data rate requires more slots then the others. The size of the first subframe is minimized if the objective is the maximization of the minimum initial data rate. In order to catch the case that links request different minimum data rate, the weight w_r is defined as

$$w_r = \frac{R_{\min,t,r}}{\sum\limits_{r \in \mathcal{R}_t} R_{\min,t,r}} \tag{4.20}$$

for all $r \in \mathcal{R}_t$. The weights are chosen such that the achieved data rate have the same ratios to each other as the minimum data rates. Then, the objective is the maximization of the minimum weighted data rate. Since the constraints that at most one resource block of the same time-frequency unit is allocated to a link and that each resource block is allocated exclusively to one link are also valid, the subproblem of the allocation of resource is given by the integer program (4.14), where the weights are defined according to (4.20). Since the weights w_r are not a function of the assignment variable $u_{r,k}$, the integer program (4.14) remains the same. Hence, the resource block allocation algorithm presented in Fig. 4.3 is applied to find a suitable allocation with a low complexity. Only the definition of the weights w_r is given by (4.20) instead of (4.13).

In order to improve the number of bits allocated to the links, the subproblem of the allocation of power and bits is solved. The subproblem is described by the same integer program as introduced for the allocation of power and bits in Section 4.3.3. According to the same motivation as given above, the minimum of the weighted initial data rate is maximized while the constraints that the power is limited and that exactly one modulation and coding scheme is used for a resource block are valid. The subproblem is described by the integer program (4.16). The algorithm illustrated in Fig. 4.4 is applied. The difference to Section 4.3.3 is that the weights w_r are defined by (4.20).

Finally, the initial data rate value $\hat{R}_{t,r}^{(1)}$ related to the first subframe allocated to RS t is found for each link (t, r).

The subproblems of the allocation of resource blocks and the allocation of power and bits are formulated for the second subframe allocated to RS t in the following. The aim is the allocation of resource blocks, power and bits such that the sum rate is maximized. Both subproblems are treated in the following according to the same structure: at first a subproblem is formulated as an integer program, then an adaptive algorithm solving the subproblem and requiring a low computational complexity is proposed.

In order to describe the allocation of resource blocks, the assignment variable $u_{r,k}$ is used again and a uniform power allocation is assumed intially. Again weights are introduced to give a general description of the objective. This time the variable $w_{r,k}$ is used. This will help in the problem formulation of the BS as introduced in Section 4.4.4 since the weights differ for all receiving UEs and for all resource blocks. For the RS, all links and all resource blocks have the same impact on the sum rate. The weight is

$$w_{r,k} = 1 \tag{4.21}$$

for each $r \in \mathcal{R}_t$ and each k. As for the integer program (4.14), the constraints given by (2.10) and (2.11) and describing that at most one resource block of the same time-frequency unit is allocated to a link and that each resource block is allocated exclusively to one link must be satisfied. Only the objective function of (4.14) is modified. The subproblem is given by

$$\max_{u_{r,k}} \sum_{r \in \mathcal{R}_t} \sum_{k=1}^{K_t} w_{r,k} u_{r,k} \epsilon_{r,k} \tag{4.22a}$$

subject to:

$$\sum_{k=(f-1)G_t+1}^{fG_t} u_{r,k} \leq 1, \qquad 1 \leq f \leq F,\ r \in \mathcal{R}_t, \tag{4.22b}$$

$$\sum_{r \in \mathcal{R}_t} u_{r,k} = 1, \qquad 1 \leq k \leq K_t, \tag{4.22c}$$

$$u_{r,k} \in \{0, 1\}, \qquad 1 < k \leq K_t,\ r \in \mathcal{R}_t. \tag{4.22d}$$

The subproblem is solved by a greedy algorithm as motivated in the following. First assume that $G_t = 1$, i.e. spatial multiplexing is neglected. Then, constraint (4.22b) is omitted since it is fulfilled for each possible allocation defined by (4.22c). In this case,

Step 1: *Initialization*

$\mathcal{K} := \{1, 2, ..., K_t\}$

$\mathcal{K}_r := \{\}$ for all $r \in \mathcal{R}_t$

Step 2: *Increase sum rate*

for $k = 1$ to K_t **do**

$r^* := \underset{r \in \mathcal{R}_t}{\operatorname{argmax}}\{\gamma_{r,k}\}$

$\mathcal{K}_{r^*} := \mathcal{K}_{r^*} \cup \{k\}$

$k \mapsto f^*$

$\gamma_{r^*,k^*} := -\infty$ for $(f^* - 1)G_t + 1 \leq k^* \leq f^* G_t$

Figure 4.5. Resource block allocation algorithm aiming at maximizing the sum rate according to the objective of (P2).

the subproblem is solved optimally if each resource block k is allocated to the best fitting link defined as the link with the largest value $\epsilon_{r,k}$ for all $r \in \mathcal{R}_t$ [LL05]. Since the allocation of two different resource blocks do not affect each other, the subproblem is solved if the best fitting link is found $K_t = F$ times. If $G_t > 1$, constraint (4.22b) must not be neglected. Then, the allocation of resource blocks defined by the same time-frequency unit f affects each other since a link must be allocated at most one resource block defined by f. A brute force solution leading to an optimum result of integer program (4.22) is to check all possible allocations of resource blocks for one time-frequency unit and to repeat this solution for all F time-frequency unit. However, the complexity of this solution increases with the number $|\mathcal{R}_t|$ of links and number of beams G_t in a time-frequency unit exponentially. In order to find an allocation of resource blocks with a lower complexity, the greedy approach is applied even for $G_t > 1$. The pseudo code of the proposed greedy algorithm is illustrated in Fig. 4.5. After the initialization of the set $\mathcal{K}$ of available resource blocks and of the set $\mathcal{K}_r$ of resource blocks allocated to link (t, r), all resource blocks are allocated successively. For each resource block, the link (t, r) providing the maximum SINR is searched. It is prevented that a link is allocated more than one resource block of the same time-frequency unit. The computational complexity required for the adaptive algorithm is given in Table 4.5.

The initial data rate value $\hat{R}_{t,r}^{(2)}$ related to the second subframe allocated to RS t is found for each link (t, r) by the algorithm depicted in Fig. 4.5. The initial data rate values are improved by the solution of the subproblem of the allocation of power and bits. This subproblem is defined by changing the objective function of the integer program

Table 4.5. Computational complexity of the resource block allocation algorithm aiming at maximizing the sum rate.

Multiplications/divisions	0
Additions/subtractions	0
Search maximum/minimum	search FG_t times in set of size $\|\mathcal{R}_t\|$

(4.16). The objective is to maximize the sum rate. The weights $w_{r,k}$ are defined for each r according to (4.21). The same constraints as in subproblem (4.16) are valid since the sum of the transmitted power is limited and a modulation and coding scheme must be allocated to each resource block. The subproblem is written as the following integer program:

$$\max_{u_{k,\epsilon}} \sum_{r\in\mathcal{R}_t} \sum_{k\in\mathcal{K}_r} \sum_{\epsilon\in\mathcal{E}} w_r u_{k,\epsilon}\epsilon \tag{4.23a}$$

subject to:

$$\sum_{k=1}^{K_t} \sum_{\epsilon\in\mathcal{E}} p_k(\epsilon) u_{k,\epsilon} \le P_t, \tag{4.23b}$$

$$\sum_{\epsilon\in\mathcal{E}} u_{k,\epsilon} = 1, \qquad 1 \le k \le K_t, \tag{4.23c}$$

$$u_{k,\epsilon} \in \{0,1\}, \qquad 1 < k \le K_t,\ \epsilon \in \mathcal{E}. \tag{4.23d}$$

Due to constraint (4.23b), the subproblem is formulated as an non-linear integer program if $G_t > 1$, where constraint (4.23b) is already discussed in Section 4.3.3. Hence, the subproblem is hard to solve. In order to find an allocation, a greedy algorithm based on the same ideas as presented in Section 4.3.3 is proposed. The proposed greedy algorithm consists of two steps and is shown in Fig. 4.6. In the first step, the bits ϵ_k allocated to resource block k are set to zero. Note that the receiver r is defined for all k uniquely. The sum p of the allocated power values is also zero. The increment $||\Delta\mathbf{p}_k||$ of power is determined for all resource blocks. In the second step, the sum p is increased as long as the limit P_t is not exceeded. The best fitting resource block k^* is searched over all resource blocks. If the limit P_t allows, additional $\Delta\epsilon$ bits are allocated to k^* and the increments are updated for all resource blocks of the same time-frequency unit denoted by f^*. The algorithm ends if the power not allocated so far is too small in order to increase the number of bits further.

If $G_t = 1$, the subproblem is given by a linear integer program solved by the algorithm optimally as shown in [BGMW07]. The computational complexity required for the adaptive algorithm is given in Table 4.6. The computational complexity includes the

Step 1: *Initialization*

$\epsilon_k := 0$ for all $k \in \mathcal{K}$

calculate $||\Delta\mathbf{p}_k||$ according (4.19) to for all k

$p := 0$

Step 2: *Increase data rates*

while $p \leq P_t$ **do**

$$k^* := \underset{k \in \bigcup_{r \in \mathcal{R}_t} \mathcal{K}_r}{\operatorname{argmin}} \{||\Delta\mathbf{p}_k||\}$$

$p := p + ||\Delta\mathbf{p}_{k^*}||$

if $p \leq P_t$, **do** $\epsilon_{k^*} := \epsilon_{k^*} + \Delta\epsilon$

$k^* \mapsto f^*$

update $||\Delta\mathbf{p}_k||$ according to (4.19) for $(f^* - 1)G_t + 1 \leq k \leq f^* G_t$

if $\epsilon_{k^*} = \max\{\mathcal{E}\}$, **do** $||\Delta\mathbf{p}_{k^*}|| := \infty$

Figure 4.6. Bit and power loading algorithm aiming at maximizing the sum rate according to the objective of (P2).

calculation of $\Delta\mathbf{p}_k$ by solving the system of linear equations given by (4.19). The worst case assumption is made that each resource block is allocated the maximum number of bits. It is assumed that a linear equation system is solved by a matrix inversion. In addition to the values given in Table 4.6, $(|\mathcal{E}| - 1)FG_t$ square roots operations are required in order to determine the 2-norm of $\Delta\mathbf{p}_k$.

Table 4.6. Computational complexity of the bit and power loading algorithm aiming at maximizing the sum rate.

Multiplications/divisions	$(\lvert\mathcal{E}\rvert - 1)FG_t(\frac{1}{2}G_t^3 + \frac{7}{2}G_t^2 + G_t)$
Additions/subtractions	$(\lvert\mathcal{E}\rvert - 1)FG_t(\frac{1}{2}G_t^3 + \frac{1}{2}G_t^2 + G_t + 1)$
Search maximum/minimum	search $(\lvert\mathcal{E}\rvert - 1)FG_t$ times in set of size FG_t

When the initial allocation is finished for both subframes allocated to RS t, the RS t knows two initial data rate values for each link (t, r), where $r \in \mathcal{R}_t$. The value $\hat{R}_{t,r}^{(1)}$ corresponds to link (t, r) and the first subframe, while the value $\hat{R}_{t,r}^{(2)}$ is corresponding to link (t, r) and the second subframe. These values are sent to the BS in order to support the allocation provided by the BS.

4.4 Initial Allocation of Resource Blocks, Power and Bits by the BS

4.4.1 Open Questions and Preliminaries

In this section, the open questions and preliminaries are given for the initial allocation of resource blocks, power and bits. For the BS $t = 0$, the questions must be answered which resource blocks are allocated to the links $(0, r)$, where $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\text{RS}}\}$, and which power is allocated to a resource block in order to apply a certain modulation and coding scheme. The BS is aware of:

- the CQI values defined as the instantaneous channel gain values and noise power values for each time-frequency unit and for each link $(0, r)$, where $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\text{RS}}\}$.
- the SINR threshold $\gamma_{\epsilon,r}$ for each number ϵ of bits, where $\epsilon \in \mathcal{E}$, and for each receiving UE or RS denoted by $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\text{RS}}\}$ in order to guarantee the bit error probability $BEP_{0,r}$ for each link $(0, r)$.
- the minimum data rate $R_{\min,0,r}$ for each link $(0, r)$ if the objective is to maximize the sum rate as claimed in problem (P2).
- the initial data rate value $\hat{R}_{(t',r)}$ for each link (t', r), where $r \in \mathcal{R}_{t'}$ and $t' \in \mathcal{R}_{\text{RS}}$ if the allocation of resource blocks, power and bits is solved by the RSs through the non-adaptive algorithm or with respect to the maximization of the minimum user rate.
- the initial data rate values $\hat{R}^{(1)}_{t',r}$ and $\hat{R}^{(2)}_{t',r}$ for each link (t', r), where $r \in \mathcal{R}_{t'}$ and $t' \in \mathcal{R}_{\text{RS}}$ if the allocation of resource blocks, power and bits is solved by the RSs with respect to the maximization of the sum rate.

The subproblem must be solved without knowing the subframe sizes in a frame.

4.4.2 Non-Adaptive Allocation of Resource Blocks, Power and Bits by the BS

In this section, the non-adaptive algorithms applied by the BS are presented. The algorithm is similar to the one of the RS presented in Section 4.3.2: Based on a uniform power allocation, resource blocks are allocated to the links by a weighted round robin algorithm which considers the bits which can be transmitted in the already allocated resource blocks. At first, the non-adaptive allocation of power and bits is treated, then the allocation of the resource blocks.

As in Section 4.3.2, the power is allocated uniformly. The power $p_{0,k}$ allocated to resource block k is given by (4.9). This leads to the SINR value $\gamma_{r,k}$ for the receiver r and the resource block k if the CQI values are taken into account. The number $\epsilon_{r,k}$ of bits is given by (4.10) if the resource block k is allocated to r.

The resource blocks are allocated to the served links by using the weighted round robin algorithm. The definition of the weight w_r of the link $(0, r)$ depends on the objective of the applied concept and on the fact if the receiving station r is a UE or a RS. If the objective of the concept is the solution of (P1) aiming at maximizing the minimum user rate, the resource blocks shall be allocated among the UEs uniformly. If the receiving station r is a UE, i.e., $r \in \mathcal{R}_0$, the weight is defined by (4.11), i.e., $w_r = 1$. Since the RS r receives the data of $|\mathcal{R}_r|$ UEs, a uniform allocation of resource blocks gives the RS more resources proportional to the number of served UEs. The weight w_r is given by

$$w_r = \frac{1}{|\mathcal{R}_r|} \tag{4.24}$$

if $r \in \mathcal{R}_{\text{RS}}$. If the objective of the concept is the solution of (P2) aiming at maximizing the sum rate, the weight w_r is defined as a function of the minimum data rate as discussed in Section 4.3.2. The minimum data rate of the UE r, where $r \in \mathcal{R}_0$, is immediately given to the BS as claimed in the initial state of the concept defined in Fig. 3.1. The minimum data rate of the RS r, where $r \in \mathcal{R}_{\text{RS}}$, is given by

$$R_{\text{min},0,r} = \sum_{r' \in \mathcal{R}_r} R_{\text{min},r,r'}. \tag{4.25}$$

The weight w_r is defined as

$$w_r = \frac{R_{\text{min},0,r}}{\sum\limits_{r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\text{RS}}\}} R_{\text{min},0,r}} \tag{4.26}$$

for all $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\text{RS}}\}$. Finally, the BS knows the initial data rate values of all links in the cell when the resource blocks, power and bits are allocated non-adaptively,.

4.4.3 Adaptive Allocation of Resource Blocks, Power and Bits by the BS Aiming at Maximizing the Minimum User Rate

In this section, the subproblem of the allocation of resource blocks, power and bits is treated for the BS if the applied concept is related to the objective of maximizing the minimum user rate defined in (P1). As introduced for the RSs in Section 4.3.3, the subproblem is divided into the allocation of resource blocks and allocation of power and bits in order to reduce the complexity. Both subproblems are treated in the following.

At first resource blocks are allocated to links in order to provide each link with suitable resource blocks. If the subproblem of the allocation of resource blocks is formulated, the difference of the BS compared to a RS is that the BS serves not only UEs, but also RSs. However, the subproblem of the RS and BS are described by the same integer program given in (4.14) since a general description of the subproblem of the allocation of resource blocks is provided in Section 4.3.3. The difference that also links to the RSs are considered is caught by a different definition of the weight w_r. If the receiving station r is a RS, the weight w_r is defined according to (4.24) in order to cover that RS r serves $|\mathcal{R}_r|$ UEs. If the receiving station r is a UE, the weight w_r is given by (4.13). Since the mathematical formulation of the subproblem of the allocation of resource blocks remains the same as in Section 4.3.3, the algorithm already introduced in Fig. 4.3 is applied to solve the subproblem.

When the allocation of resource blocks is solved, the initial data rate values are increased by solving the allocation of power and bits adaptively. The subproblem of the allocation of power and bits is also represented by the same integer program (4.16) as introduced in Section 4.3.3, but the weight w_r is defined by (4.24) and (4.13) if the receiving station is a RS and a UE, respectively. The allocation of power and bits is performed by the algorithm illustrated in Fig. 4.4.

4.4.4 Adaptive Allocation of Resource Blocks, Power and Bits by the BS Aiming at Maximizing the Sum Rate

In this section, the subproblem of the allocation of resource blocks, power and bits is treated for the BS if the objective of the applied concept is the maximization of the sum rate as defined by (P2). Simultaneously, a minimum data rate is guaranteed for each link. The BS pursues the same strategy as the RS and as stated and motivated in Section 4.3.4. The BS achieves its objective in two steps. At first, the aim is to guarantee the minimum data rate for each link $(0, r)$, where $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}$. This aim is addressed in a first subframe assigned to the BS. Then, the aim is the maximization of the sum rate as addressed in a second subframe. As motivated in Section 4.3.3, the subproblem of allocating resource blocks, power and bits is split for each subframe in the two subproblems of allocating resource blocks and of allocating power and bits.

The formulations of the subproblems of allocating resource blocks and of allocating power and bits and the applied algorithms concerning the first subframe allocated to the BS yield from an straightforward extension of the considerations given in Section 4.3.4 and Section 4.4.3. As introduced in Section 4.3.4, the objective of the allocation related to the first subframe is to achieve the minimum data rate for each link. This objective is covered by the maximization of the minimum initial data rate. As derived in Section 4.4.3, the corresponding subproblems of allocating resource blocks and of allocating power and bits are described by the integer programs (4.14) and (4.16), respectively. The weight w_r is defined by (4.26) for all $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}$. If the receiving station r is a RS, the minimum data rate $R_{\min,0,r}$ results from (4.25). The algorithms illustrated in Fig. 4.3 and Fig. 4.4 are applied to solve both subproblems.

In the second subframe, the objective is the maximization of the sum rate. Due to the objective, the resource blocks, power and bits shall be allocated to the BS-to-UE and BS-to-RS links promising the maximum user rate. However, the user rate of the direct connections and two-hop connections cannot be determined since the user rate is a function of the subframe sizes which are not determined yet. Assume that the subframe size of the second subframe allocated to the BS is denoted as S_2. If $\epsilon_{r,k}$ bits are carried by the slot of a resource block allocated to a BS-to-UE link, its impact on the sum rate is $\epsilon_{r,k}S_2$ bits. If $\epsilon_{r,k}$ bits are allocated to a BS-to-RS link for one slot, its impact is only $\epsilon_{r,k}S_2$ bits, if these $\epsilon_{r,k}S_2$ bits are forwarded by the RS. Hence, the RS must be allocated slots. These slots are not available for the BS and reduce the value of S_2 when the allocation of slots is solved. A simple solution could be the neglection of the BS-to-RS links, but this is rather inefficient if a two-hop connection exist in the cell which could achieve the highest user rate. In this thesis, it is proposed to introduce a penalty for the BS-to-RS links. Using the penalty allows a comparison of the BS-to-UE link and a BS-to-RS link when resource blocks or power and bits are allocated. The penalty is used as a weight when the subproblems of the allocation of resource blocks and the allocation of power and bits are formulated. A definition of the weight is required for each resource block k and each receiving station $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}$. In conformance with the previous sections, the variable $w_{r,k}$ is used for the weight corresponding to link $(0,r)$ and resource block k. At first, the weight is derived, then the allocation of resource blocks is treated and finally, the allocation of power and bits is treated.

If a resource block is allocated to a UE, the bits allocated to the resource block have a full impact on the sum rate and are not subject to the transmission of an RS. This is modelled by the weight defined as

$$w_{r,k} = 1 \tag{4.27}$$

for $r \in \mathcal{R}_0$ and for each k. If the receiving station r is a RS, a weight smaller than one is introduced in order to model that slots are required for the transmission of the RSs. The weight is derived as follows: If one resource block is allocated to the link $(0, r)$, a resource block is required for the link (r, r'), where $r \in \mathcal{R}_{\text{RS}}$ and $r' \in \mathcal{R}_r$. The BS knows the initial data rate value of the second subframe allocated to RS r. This value is denoted by $\hat{R}^{(2)}_{r,r'}$. However, the exact allocation of resource blocks, power and bits on the link (r, r') and the number of slots per subframe are missing. Hence, two assumptions are made: one for the number of bits per resource block allocated to the RS-to-UE link and one for the subframe sizes.

The number of bits per resource block are estimated by an average value. If one slot is allocated to the RS r, the value

$$\epsilon_r = \sum_{r' \in \mathcal{R}_r} \hat{R}^{(2)}_{r,r'} / K_r. \tag{4.28}$$

represents the bits which can be forwarded by the RS in a slot of a resource block averaged over all RS-to-UE links. The value ϵ_r is used as the first assumption: Whenever a resource block is allocated by the BS to the link $(0, r)$, it is assumed that the RS forwards ϵ_r bits in a resource block and slot.

The second assumption is related to the allocation of slots to the BS and RS where only the slots are considered which are used for maximizing the sum rate. Assume for a moment that S slots are shared by the BS and a RS for a single two-hop connection. The slots used for the BS-to-RS link and for the RS-to-UE link are denoted by S_{BS} and S_{RS}, respectively. In total the BS and RS shall share the S slots such that

$$S_{\text{BS}} + S_{\text{RS}} = S \tag{4.29}$$

holds, where $S_{\text{BS}} \geq 0$ and $S_{\text{RS}} \geq 0$. If $\epsilon_{r,k}$ are transmitted on the link $(0, r)$ in a slot of resource block k and if ϵ_r bits are forwarded by RS r per slot, the slots are shared such that the RS receives as much as it can forward if

$$S_{\text{BS}} \epsilon_{r,k} = S_{\text{RS}} \epsilon_r \tag{4.30}$$

holds. Using (4.29) and (4.30), the slots used for the BS-to-RS link and related to the total number of slots are

$$\frac{S_{\text{BS}}}{S} = \frac{\epsilon_r}{\epsilon_{r,k} + \epsilon_r} < 1. \tag{4.31}$$

The ratio represents the fraction in which information is sent by the source while the fraction $\frac{S_{\text{RS}}}{S} = 1 - \frac{S_{\text{BS}}}{S}$ is representing the fraction of repetition. If resource block k is loaded with $\epsilon_{r,k}$ bits and if resource block k is allocated to RS r, the value $\frac{S_{\text{BS}}}{S}$ serves as the estimate how many bits per slot have an impact on the user rate. Then, the weight $w_{r,k}$ is defined as

$$w_{r,k} = \frac{\epsilon_r}{\epsilon_{r,k} + \epsilon_r} \tag{4.32}$$

if $r \in \mathcal{R}_{\text{RS}}$. Since the weight is smaller than one, resource block k allocated to RS r has not the full impact of $\epsilon_{r,k}$ bits on the sum rate, but its impact is only $w_{r,k}\epsilon_{r,k}$ bits.

The usage of the weights allows that the subproblems of allocating resource blocks and of allocating power and bits are described by the same integer programs as introduced for an RS in Section 4.3.4 if the allocations are treated for the second subframe assigned to an RS. The subproblem of the allocation of resource blocks combined with the objective that the sum rate is maximized is the same as defined in (4.22) but the weights are given by (4.27) and (4.32) for the BS. The subproblem is solved with a modified version of the algorithm illustrated in Fig. 4.5. The operation

$$r^* = \underset{r \in \mathcal{R}_t}{\text{argmax}}\{\gamma_{r,k}\} \tag{4.33}$$

is substituted by

$$r^* = \underset{r \in \mathcal{R}_t}{\text{argmax}}\{w_{r,k}\epsilon_{r,k}\} \tag{4.34}$$

with respect to the definition of the penalty. The computational complexity required for the adaptive algorithm is given in Table 4.7. The computational complexity includes the calculation of the penalty.

Table 4.7. Computational complexity of the resource block allocation algorithm aiming at maximizing the sum rate.

Multiplications/divisions	$2\lvert\mathcal{R}_{\text{RS}}\rvert FG_0$
Additions/subtractions	$\lvert\mathcal{R}_{\text{RS}}\rvert FG_0$
Search maximum/minimum	search FG_0 times in set of size $\lvert\mathcal{R}_{\text{RS}}\rvert + \lvert\mathcal{R}_0\rvert$

The subproblem how the BS allocates power and bits is also described by the same integer program as the corresponding subproblem of a RS and is defined in Section 4.3.4. The integer program is given for the BS and RS in (4.23). The difference that the BS serves not only UEs but also RSs is catch by defining the weights according to (4.27) and (4.32). The subproblem is solved by a modified version of the algorithm defined in Fig. 4.6. The operation

$$k^* = \underset{k \in \bigcup_{r \in \mathcal{R}_t} \mathcal{K}_r}{\text{argmin}} \{||\Delta \mathbf{p}_k||\} \tag{4.35}$$

is replaced by

$$k^* = \underset{k \in \bigcup_{r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}} \mathcal{K}_r}{\operatorname{argmin}} \left\{ \frac{||\Delta \mathbf{p}_k||}{w_{r,k} \epsilon_{r,k}} \right\}. \tag{4.36}$$

Equation (4.36) contains a normalization of the require increment of power to the weighted number of bits. The normalization is introduced in order to penalize an allocation to a BS-to-RS link. Based on the worst case assumption that the maximum number of bits is allocated, the computational complexity required for the adaptive algorithm is given in Table 4.8. In addition to the values given in Table 4.8, $(|\mathcal{E}|-1)FG_t$ square roots operations are required in order to determine the 2-norm of $\Delta \mathbf{p}_k$.

Table 4.8. Computational complexity of the bit and power loading algorithm aiming at maximizing the sum rate.

Multiplications/divisions	$(\lvert\mathcal{E}\rvert-1)FG_0(\frac{1}{2}G_0^3 + \frac{7}{2}G_0^2 + G_0 + 2)$
Additions/subtractions	$(\lvert\mathcal{E}\rvert-1)FG_0(\frac{1}{2}G_0^3 + \frac{1}{2}G_0^2 + G_0 + 1)$
Search maximum/minimum	search $(\lvert\mathcal{E}\rvert-1)FG_0$ times in set of size FG_0

4.5 Allocation of Slots to Access Points

4.5.1 Open Questions and Preliminaries

The open question must be answered how many slots are allocated to the APs. The question is answered when the BS finished its initial allocation of resource blocks, power and bits. The BS is aware of:

- the minimum data rate $R_{\min,t,r}$ for each link in the cell if the objective is to maximize the sum rate as claimed in problem (P2).
- the initial data rate value $\hat{R}_{(t,r)}$ for each link in the cell if the allocation of resource blocks, power and bits is solved by the RSs through the non-adaptive algorithm or with respect to the maximization of the minimum user rate.
- the initial data rate values $\hat{R}_{t,r}^{(1)}$ and $\hat{R}_{t,r}^{(2)}$ for each link in the cell if the allocation of resource blocks, power and bits is solved by the RSs with respect to the maximization of the sum rate.

4.5.2 Adaptive Allocation of Slots to Access Points Aiming at Maximizing the Minimum User Rate

In this section the subproblem of the allocation of slots is treated if the objective of the resource allocation concept is the maximization of the minimum user rate as defined in (P1). At first, the formulation of the subproblem as an integer program is derived. Then, the problem is rewritten in order to find an allocation of slots by solving a system of linear equations.

According to Section 4.3.3 and Section 4.4.3, exactly one subframe is assigned to each AP. Without loss of generality, it is assumed that the subframes are ordered successively to the index t of the transmitting APs, i.e., the subframe n, where $n = 1, 2, ...N_{\mathrm{SF}}$, is allocated to t if $n = t + 1$. The first subframe is allocated to the BS. The subproblem of the allocation of slots is formulated based on the relationship between the initial data rate values and the user rate. If the UE r uses a direct connection, i.e., $r \in \mathcal{R}_0$, the user rate is a function of the initial data rate $\hat{R}_{0,r}$ and the subframe size S_0 of the first subframe. The user rate is

$$R_{0,r} = S_0 \hat{R}_{0,r} \tag{4.37}$$

since the initial allocation of resource blocks, power and bit is valid for all slots allocated to a subframe. If the UE r uses a two-hop connection via the RS t, i.e., $r \in \mathcal{R}_t$ and $t \in \mathcal{R}_{\mathrm{RS}}$, the user rate depends on the initial data rate values $\hat{R}_{0,t}$ and $\hat{R}_{t,r}$ and the subframe sizes S_0 and S_t, where S_t represents the size of the subframe allocated to RS t. Since the RS receives the bits of all $|\mathcal{R}_t|$ UEs, the data rate $\hat{R}_{0,t}$ must be demultiplexed. If the objective is to maximize the minimum user rate, the initial data rate $\hat{R}_{0,t,r}$ finally addressed to r and sent by the BS to RS t is chosen such that no UE is favored. Then, the initial data rate value $\hat{R}_{0,t,r}$ is

$$\hat{R}_{0,t,r} = \frac{\hat{R}_{0,t}}{|\mathcal{R}_t|}. \tag{4.38}$$

Using (2.23), the data rate of the two-hop connection is

$$R_{0,r} = \min \left\{ S_0 \hat{R}_{0,t,r}; S_t \hat{R}_{t,r} \right\} = \min \left\{ S_0 \frac{\hat{R}_{0,t}}{|\mathcal{R}_t|}; S_t \hat{R}_{t,r} \right\}. \tag{4.39}$$

The subproblem of the allocation of slots is written as designing the subframe sizes $S_0, S_1, ...S_{N_{\mathrm{RS}}}$ such that the minimum user rate is maximized. Taking into account the constraint related to the length S of the frame and given by (2.1), the subproblem is represented by the following integer program:

$$\max_{S_0, S_1, ...S_{N_{\mathrm{RS}}}} \quad \min_{r \in \bigcup_{t=1}^{N_{\mathrm{RS}}} \mathcal{R}_t} \quad R_{0,r} \tag{4.40a}$$

subject to:

$$\sum_{t=0}^{N_{\mathrm{RS}}} S_t = S, \tag{4.40b}$$

$$S_t \in \{0, 1, ..., S\}, \qquad \text{for all } t. \tag{4.40c}$$

In order to solve the integer program, the integer program is rewritten. Two changes are introduced.

1. Instead of accounting for all links, only one link of the subframe allocated to RS t is considered in the integer program. For the subframe of RS t, only the link $(t, r_{\min,t})$, where
$$r_{\min,t} = \underset{r \in \mathcal{R}_t}{\operatorname{argmin}}\{\hat{R}_{t,r}\}, \tag{4.41}$$
is considered since all the other links do not affect the minimum user rate.

2. An equivalent formulation of the objective is introduced. Only the weighted data rate of the weakest BS-to-UE link is maximized while the other weighted data rates are ensured to be larger than the one of the weakest link. The weights related to the BS-to-UE and BS-to-RS links are defined as
$$w_r = \left\{ \begin{array}{ll} 1 & \text{if } r \in \mathcal{R}_0 \\ \frac{1}{|\mathcal{R}_r|} & \text{if } r \in \mathcal{R}_{\mathrm{RS}}. \end{array} \right\} \tag{4.42}$$
in order to model that the RSs serve multiple UEs. The weakest BS-to-UE link is $(0, r_{\min,0})$, where
$$r_{\min,0} = \underset{r \in \mathcal{R}_0}{\operatorname{argmin}}\{w_r \hat{R}_{0,r}\}. \tag{4.43}$$
It is claimed that all the other links, achieve weighted data rates larger than or equal to the weighted data rate of the weakest link.

Integer program (4.40)is rewritten as

$$\max_{S_0, S_1, \ldots S_{N_{\mathrm{RS}}}} \quad \hat{R}_{0,r_{\min,0}} S_0 \tag{4.44a}$$

subject to:

$$\sum_{t=0}^{N_{\mathrm{RS}}} S_t = S, \tag{4.44b}$$

$$S_t \in \{0, 1, ..., S\}, \qquad \text{for all } t, \tag{4.44c}$$

$$S_0 \hat{R}_{0,t} \geq |\mathcal{R}_t| S_t \hat{R}_{t,r_{\min,t}}, \qquad \text{for all } t \in \mathcal{R}_{\mathrm{RS}}. \tag{4.44d}$$

Constraint (4.44d) ensures that the bits transmitted to the RS t given by the left-hand side of the inequation are larger than the multiple of the bits received by the weakest link of RS t. The allocation of slots is solved most efficiently if (4.44d) is fulfilled with equality. If so, equation (4.44d) and (4.44b) form a set of $N_{\mathrm{RS}}+1$ linearly independent equations. The solution of the system of linear equations yields

$$S_0 = \frac{S \prod_{t'=1}^{N_{\mathrm{RS}}} \hat{R}_{t',r_{\min,t'}}}{\prod_{t'=1}^{N_{\mathrm{RS}}} \hat{R}_{t',r_{\min,t'}} + \sum_{t'=1}^{N_{\mathrm{RS}}} w_{t'} \hat{R}_{0,t'} \prod_{\substack{t''=1\\t''\neq t'}}^{N_{\mathrm{RS}}} \hat{R}_{t'',r_{\min,t''}}} \tag{4.45}$$

and

$$S_t = \frac{S w_t \hat{R}_{0,t} \prod_{\substack{t'=1\\t'\neq t}}^{N_{\mathrm{RS}}} \hat{R}_{t',r_{\min,t'}}}{\prod_{t'=1}^{N_{\mathrm{RS}}} \hat{R}_{t',r_{\min,t'}} + \sum_{t'=1}^{N_{\mathrm{RS}}} w_{t'} \hat{R}_{0,t'} \prod_{\substack{t''=1\\t''\neq t'}}^{N_{\mathrm{RS}}} \hat{R}_{t'',r_{\min,t''}}}. \tag{4.46}$$

if $t \in \mathcal{R}_{\mathrm{RS}}$. Note that the optimum solution of (4.40) is found by (4.45) and (4.45) if $S_0, S_1, ...S_{N_{\mathrm{RS}}}$ are integer values. Since subproblem (4.40) is a max-min problem, the number of bits transmitted in the subframes are balanced in the optimum solution. If the equation (4.45) and (4.45) do not yield integer values, only the values leading to an upper bound of subproblem (4.40) are found, but the actual solution is not found. The optimum solution of subproblem (4.40) is found by a local search. The values of $S_0, S_1, ...S_{N_{\mathrm{RS}}}$ are rounded to the next lower or larger integer value. In total, $2^{N_{\mathrm{RS}}+1}$ combinations of rounding to the next lower or larger integer value exist for the $N_{\mathrm{RS}}+1$ subframes. The optimum is found by testing all these combinations which lead to a frame size of S slots. However, the influence of a single slot allocated to a subframe or not vanishes if the frame consists of a large number of slots. Hence, a more practical approach which saves the computational complexity required for a local search is to choose an arbitrary combination which leads to a frame size of S slots.

4.5.3 Adaptive Allocation of Slots to Access Points Aiming at Maximizing the Sum Rate

In this section, the allocation of slots is treated if the objective of the resource allocation concept is the maximization of the sum rate as defined in (P2). According to Section 4.3.4 and Section 4.4.4, two subframes are assigned to each AP. In the first type of subframes, resource blocks, power and bit are allocated with the aim to achieve the minimum data rate. These type is considered in this section at first. The second type of subframes serves for the maximization of the sum rate and is treated in the second part of this section.

The size of the subframes is derived as follows: Similar to the previous section, the assumption is made without loss of generality that the subframes are allocated to the APs in the successive order according to the variable t. The value of the index n, where $n = 1, 2, ...N_{\mathrm{SF}}$, is mapped to the AP t. Subframe $2t+1$ and $2t+2$ are assigned to AP t. The size of the subframe $2t+1$ assigned to AP t is chosen such that the minimum data rate is achieved by each links (t, r), where $r \in \mathcal{R}_t$. Since the data rate of each link served in the subframe increases with the number of slots linearly, the weakest link determines the number of required slots. The receiver of the weakest link is found by (4.41) and (4.43) if the AP t is a RS and a BS, respectively. Then, the size of subframe $2t+1$ is given by

$$S_{2t+1} \geq \frac{R_{\mathrm{min},t,r_{\mathrm{min},t}}}{\hat{R}_{t,r_{\mathrm{min},t}}}. \tag{4.47}$$

It is most efficient if (4.47) is fulfilled with equality. If the ratio of the right hand side is not an integer number, the next larger integer number must be chosen. If each AP is allocated so many slots that the minimum data rate is achieved on each link, the number of remaining slots is

$$S_{\mathrm{max}} = S - \sum_{t=0}^{N_{\mathrm{RS}}} S_{2t+1}. \tag{4.48}$$

These S_{max} are shared by the APs with the aim to maximize the sum rate. Note, if $S_{\mathrm{max}} < 0$, not all links can achieve their demanded minimum data rate. If this problem occurs, a solution is to drop the UEs served by at least one RS or to accept a reduction of the minimum data rate for the UEs served by at least one RS. It depends on the offered service which solution is chosen and how a solution looks like in detail. If $S_{\mathrm{max}} = 0$, the allocation of subframes is solved since remaining slots do not exist.

In the following, the case is consider where $S_{\mathrm{max}} > 0$ and the sizes of the remaining subframes are derived. The problem must be solved how these $S_{\mathrm{max}} > 0$ slots are allocated to the APs such that the sum rate is maximized. Since the slots allocated to an AP are grouped in a subframe, the problem reduces to the problem to find the integer values of slots allocated to the APs. The problem is formulated as a sequential decision process in the following in order to motivate its solution.

The slots are allocated by S_{max} decisions. Each decision is made at a time instant s, where $s = 1, 2, ...S_{\mathrm{max}}$. The number of slots allocated to the APs after s decisions are denoted by the vector $\mathbf{s}_s$, where $\mathbf{s}_s \in \mathbb{N}_0^{(N_{\mathrm{RS}}+1)\times 1}$. Initially, none of the S_{max} slots are allocated to any AP. Since a RS can only forward when it has received data from the BS, the first decision is that the first slot is allocated to the BS and

$$\mathbf{s}_1 = (1, 0, 0, ...0)^{\mathrm{T}} \tag{4.49}$$

holds. In the decision s, an AP is chosen. This is denoted by a decision variable $\mathbf{x}_s$ consisting of $N_{\mathrm{RS}} + 1$ elements and defined as

$$\mathbf{x}_s \in \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdots \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right\}. \tag{4.50}$$

Since decisions are made sequentially, the number of allocated slots fulfills the constraint

$$\mathbf{s}_s = \mathbf{s}_{s-1} + \mathbf{x}_s. \tag{4.51}$$

If a slot is allocated to the AP t, the AP can sent at most

$$\hat{R}_t^{(2)} = \sum_{r \in \mathcal{R}_t} \hat{R}_{t,r}^{(2)} \tag{4.52}$$

bits per slot, where $\hat{R}_{t,r}^{(2)}$ is the initial data rate value of the link (t, r) in the second subframe allocated to AP t. If the AP t is the BS, these bits per slot can always be sent and increase the sum rate. If the AP t is a RS, bits must be available at the RS for transmission. The number of bits which are available for the RS, are given by the difference between the bits received from the BS and the bits already considered for transmission. The bits received from the BS at decision s are denoted by $\hat{R}_{0,t}^{(2)} s_{(s-1),1}$, where $s_{(s-1),t+1}$ represents the t-th+1 element of the vector $\mathbf{s}_{s-1}$. The bits already considered for transmission are given by $\hat{R}_t^{(2)} s_{(s-1),t}$. Note, that both values are a function of $s - 1$ and not of s. The bits which can be transmitted by the APs if a slot is allocated at decision s are represented by

$$\mathbf{f}_s = \begin{pmatrix} \hat{R}_0^{(2)} \\ \min\left\{ \hat{R}_{0,1}^{(2)} s_{(s-1),0} - \hat{R}_1^{(2)} s_{(s-1),1}, \hat{R}_1^{(2)} \right\} \\ \min\left\{ \hat{R}_{0,2}^{(2)} s_{(s-1),0} - \hat{R}_2^{(2)} s_{(s-1),2}, \hat{R}_2^{(2)} \right\} \\ \vdots \\ \min\left\{ \hat{R}_{0,N_{\mathrm{RS}}}^{(2)} s_{(s-1),0} - \hat{R}_{N_{\mathrm{RS}}}^{(2)} s_{(s-1),N_{\mathrm{RS}}}; \hat{R}_t^{(2)} \right\} \end{pmatrix}, \tag{4.53}$$

where the first element of $\mathbf{f}_s$ corresponds to the BS and the other elements to the RSs. The subproblem of the allocation of slots such that the sum rate is maximized is written as

$$\max_{\mathbf{x}_2, \mathbf{x}_3, \ldots \mathbf{x}_{S_{\max}}} \sum_{s=2}^{S_{\max}} \mathbf{f}_s^{\mathrm{T}} \mathbf{x}_s \tag{4.54a}$$

Step 1: *Find non-empty subframes*

$\mathcal{T} := \{0\}$

for $t = 1$ to $|\mathcal{R}_{\mathrm{RS}}|$ **do**

if $\hat{R}_t^{(2)} > \hat{R}_0^{(2)}$, $\mathcal{T} := \mathcal{T} \cup \{t\}$

if $\mathcal{T} = \{0\}$ **do** allocate all $S_{\max}$ slots to BS and break algorithm

Step 2: *Initialization*

$\mathbf{s}_1 := (1, 0, 0, ...0)^{\mathrm{T}}$ with dimension $|\mathcal{T}| \times 1$

initialize $\mathbf{f}_2$ according to (4.53)

Step 3: *Build subframes*

for $s = 2$ to $S_{\max}$ **do**

$t^* := \underset{t \in \mathcal{T}}{\operatorname{argmax}} \{\mathbf{f}_s\}$

$s_{s,t^*} := s_{(s-1),t^*} + 1$

$s_{s,t'} := s_{(s-1),t'}$ for $t' \neq t^*$

update $\mathbf{f}_{s+1}$ according to (4.53)

Figure 4.7. Allocation of slots aiming at maximizing the sum rate according to the objective of (P2).

subject to:

$$\mathbf{s}_s = \mathbf{s}_{s-1} + \mathbf{x}_s, \qquad 2 \leq s \leq S_{\max} \tag{4.54b}$$

$$\mathbf{s}_1 = (1, 0, 0, ...0)^{\mathrm{T}}. \tag{4.54c}$$

In other words, the subproblem is to find these $S_{\max} - 1$ decision variables which lead to a maximum sum rate. The subproblem is restricted to the constraints that the decisions are made sequentially and that a slot is allocated to the BS in the first decision. The integer program (4.54) has the structure of a so called sequential decision process, e.g., treated in [NW99]. Its solution is found by dynamic programming. Dynamic programming is a set of algorithms which solve optimization problems related to sequential decision processes by decomposing the problem in smaller ones. These smaller problems are solved sequentially. The solution of such a small problem depends only on the solution of the one solved previously. A famous example of a dynamic programming algorithm is the Viterbi algorithm [For73].

In order to solve the subproblem represented by (4.54), a dynamic programming algorithm is proposed. The algorithm is illustrated in Fig. 4.7 and is structured in three steps. In the first step, the number of considered APs is grouped in the set $\mathcal{T}$. The BS must always be considered. If the RS t cannot forward more bits per slot to the UEs than the BS, i.e., $\hat{R}_t^{(2)} \leq \hat{R}_0^{(2)}$, it is always preferable to allocate the slots to the BS and not to RS t. Hence, the RS t is only an element of $\mathcal{T}$, if $\hat{R}_t^{(2)} > \hat{R}_0^{(2)}$. If no RS is element of $\mathcal{T}$, the algorithm can be stopped because all $S_{\max}$ slots are allocated to the BS. In the second step, the vectors $\mathbf{s}_1$ and $\mathbf{f}_2$ are initialized. In the third step, the subframe sizes are determined in $S_{\max} - 1$ iterations. Within an iteration, the AP fitting best to $\mathbf{f}_s$ is chosen and the number of slots allocated to the APs and the function $\mathbf{f}_s$ are updated. The optimum solution of (4.54) is found in $S_{\max}$ steps. Note that the search of the best fitting AP is simplified if not all RSs are within the set $\mathcal{T}$.

The number of slots allocated to AP t in subframe $2t + 2$ is given by the t-th+1 element of the vector $\mathbf{s}_{S_{\max}}$. The sequential order in which the slots are allocated to the AP does not matter since only the subframe size must be determined.

The computational complexity required for the adaptive algorithm is given in Table 4.9. The expressions are determined based on the assumption that (4.53) is re-calculated for each state completely and intermediate results are not stored. Additionally, it is assumed that all N_{RS} RSs are considered in the set $\mathcal{T}$.

Table 4.9. Computational complexity of the allocation of slots aiming at maximizing the sum rate.

Multiplications/divisions	$2N_{\text{RS}}(S_{\max} - 1)$
Additions/subtractions	$(N_{\text{RS}} + 1)(S_{\max} - 1)$
Search maximum/minimum	search $S_{\max} - 1$ times in set of size $N_{\text{RS}} + 1$, search $(S_{\max} - 1)N_{\text{RS}}$ times in set of size 2

4.6 Final Allocation of Resource Blocks, Power and Bits

In this section, it is described how the allocation of resource blocks, power and bits is finalized in order to achieve the allocation of resources.

When the allocation of slots is solved by the BS, the BS is aware of the subframe sizes. The solutions related to the allocation of resource blocks and to the allocation of power and bits are actually considered in Section 4.3 and Section 4.4 for one slot. However, the solutions are valid for all slots. The user rate of a direct connection is given by (2.22) and the one of a two-hop connection by (2.23). Since the number of slots must be an integer number, the data rate of the BS-to-RS link and of the RS-to-UE link are not always balanced for a two-hop connection perfectly. For a transmission from the BS to RS r' and from RS $t' = r'$ to UE r, the difference denoted as Δ is

$$\Delta = |R_{0,r',r} - R_{t',r}| \tag{4.55}$$

and is given in bits per frame. If $\Delta > 0$, the BS can send more data to the RS than the RS is able to forward. In order to balance the data rates, the BS reduces the data rate by selecting more robust modulation and coding schemes for single resource blocks as long as Δ does not become negative. The allocation of power keeps unchanged. Hence, the bit error probability is improved on the link. If $\Delta \neq 0$ although the modulation and coding schemes are changed, zero padding is applied additionally to find a perfectly balanced state. If $\Delta < 0$, the RS can forward more data then it receives. Then, the RS chooses more robust modulation and coding schemes and applies zero padding if required.

4.7 Summary

In this chapter, the subproblems of the design of grids of beams, the allocation of resource blocks, the allocation of power and bits and the allocation of slots are introduced for the distributed concept for orthogonal medium access. Novel algorithms enabling an adaptive allocation at a low complexity are presented. Except for the allocation of slots, non-adaptive algorithms are presented, too. Although, the adaptive algorithms are expected to perform better in terms of the objective defined in (P1) and (P2), the adaptive algorithms are only applied, if the computational complexity is available. If this is not the case the distributed concept for orthogonal medium access allows to replace an adaptive algorithm by the corresponding non-adaptive algorithms. The allocation of slots is excluded since a non-adaptive algorithms affects the allocation of resource blocks, power and bits. This is treated in the distributed concept for reuse medium access and addressed in the next chapter.

Chapter 5
Distributed Concept for Reuse Medium Access

5.1 Introduction

In this chapter, the subproblems which must be solved in the distributed concept for reuse medium access are formulated. The organization of the chapter is according to the successive order of the subproblems defined by the distributed concept for reuse medium access. Hence, Fig. 3.2 acts as a roadmap of this chapter. In Sections 5.2, the subproblem of the allocation of slots is treated. A non-adaptive algorithm is proposed, where slots are allocated to an AP depending on the probability that UEs are expected in the coverage area of an AP. In Sections 5.3 to 5.5, the subproblems of the design of grids of beams, the initial allocation of resource blocks solved by the RS and the allocation of resource blocks solved by the BS are addressed successively. For each of the three subproblems, the open questions and the preliminaries related to the subproblem are formulated, a non-adaptive algorithm is proposed and an adaptive algorithm is derived from the description of the subproblem as a linear or non-linear integer program. The computational complexity of the non-adaptive algorithms is negligible. Each adaptive algorithm finds a suboptimum solution of the corresponding integer program, but requires a low computational complexity compared to an algorithm leading to an optimum solution like an exhaustive search algorithm, dynamic programming algorithm or branch-and-bound algorithm. In Section 5.6, the final allocation of resource blocks and bits is treated in order to balance the data rate of the BS-to-RS links and the data rate of the RS-to-UE links. Section 5.7 summarizes the main contributions of this chapter.

5.2 Non-Adaptive Allocation of Slots to Access Points

In this section, the non-adaptive algorithm which allocates slots to the APs is presented. Each AP is allocated one subframe. The size of the subframe is determined once and kept for all frames. An experimental approach solving the subproblem is that various combinations are tested for the subframe sizes among several frames when the operation of the relay network starts. Then, the best fitting one is chosen and kept. This approach is only suitable if the number of slots per frame is small and the number of tested combinations are small. In order to find a solution for large frames, a simple non-adaptive algorithm is motivated and introduced. The size is determined based on the PDF $\rho(x, y)$ representing the distribution of the UEs in the cell, the reuse groups $\mathcal{T}_n$ for each subframe n, where $n = 1, 2, ...N_{\mathrm{SF}}$ and the coverage area A_t for each AP t. Instantaneous information like the current assignment of UEs to APs, CQI values or initial data rate values are not considered since the allocation is valid for each frame. The objective is that the slots are allocated to the APs depending on the probability that UEs are expected in the coverage area of an AP. Slots are allocated uniformly among the number of expected UEs. The objective is chosen with respect to the objective of maximizing the minimum user rate as defined in (P1) since a uniform allocation is beneficial in order to maximize the minimum user rate if instantaneous information are not available. The objective is also chosen with respect to the maximization of the sum rate as defined in (P2) since a uniform allocation is beneficial in order to achieve the minimum user rate for all UEs.

Without loss of generality, it is assumed that the subframes are ordered successively to the index of the reuse groups, i.e., the subframe n is allocated to reuse group $\mathcal{T}_n$. The first reuse group $\mathcal{T}_1$ consists only of the BS and the BS is not an element of a second reuse group. The case, in which the BS is an element of several reuse groups and RSs are transmitting simultaneously, is analyzed in [SAY06] , where further details can be found for this case. For simplicity, it is assumed further that all RSs apply the same number of beams denoted by G_{RS}.

At first, only the reuse groups consisting of RSs are considered. A uniform allocation among the served UEs is given if RSs which are expected to serve more UEs are allocated more slots. The expectation about the number of UEs is represented by the probability Pr_t that a user is in the coverage area A_t of RS t, where the coverage are A_t is introduced in Section 4.2.3. The probability Pr_t is given by

$$Pr_t = \int\int_{A_t} \rho(x, y)\ dxdy. \tag{5.1}$$

If several RSs transmit in the same reuse group, the same subframe size is allocated to all RSs in the reuse group. As stated in Section 2.3, the user distribution must be considered in the design of the reuse groups and it is optimum if each RS t in the reuse group has the same probability Pr_t. If this is not the case, the one with the largest Pr_t is chosen. Its probability is denoted by Pr_n in the following since it is representative for the reuse group $\mathcal{T}_n$. The other RSs of $\mathcal{T}_n$ do not affect the subframe size since a smaller amount of UEs is expected for them. The ratio of the size of subframe n and n', where $n' = 2, 3, ...N_{\mathrm{SF}}$ and $n' \neq n$, shall fulfil

$$\frac{S_{n'}}{S_n} = \frac{Pr_{n'}}{Pr_n} \tag{5.2}$$

in order to achieve a uniform allocation.

Now the reuse group containing the BS and denoted by $\mathcal{T}_1$ is considered. The subframe size S_1 is related to the subframe sizes S_n, where n represents a subframe in which RSs transmit, i.e., $n = 2, 3, ...N_{\mathrm{SF}}$. The subframe sizes are weighted by the number of beams applied by the APs and denoted by G_0 and G_{RS} since the time-frequency units of a subframe are reused in space. With respect to the fact that the BS provides the bits per frame of all UEs in the cell, the probability Pr_1 is set to one. The probability Pr_1 is related to the probability representing the reuse group $\mathcal{T}_n$. Then, the ratio

$$\frac{Pr_n}{Pr_1} = \frac{G_{\mathrm{RS}} S_n}{G_0 S_1} \tag{5.3}$$

shall be fulfilled. If all combinations of reuse groups are considered, equation (5.2) and (5.3) lead to $N_{\mathrm{SF}} - 1$ linearly independent equations. Additionally, the sum of the subframe sizes is limited, i.e., constraint (2.1) must be fulfilled. Hence, a system of N_{SF} linear equations must be solved. Its solution yields that the size of the first subframe is

$$S_1 = \frac{S \cdot G_{\mathrm{RS}} \cdot Pr_1}{G_{\mathrm{RS}} Pr_1 + G_0 \sum_{n=2}^{N_{\mathrm{SF}}} Pr_n}. \tag{5.4}$$

The size of subframe n, where $n = 2, 3, ...N_{\mathrm{SF}}$, is

$$S_n = \frac{S \cdot G_0 \cdot Pr_n}{G_{\mathrm{RS}} Pr_1 + G_0 \sum_{n=2}^{N_{\mathrm{SF}}} Pr_n}. \tag{5.5}$$

However, only an integer value of slots can be assigned to a subframe and a rounding is possibly required.

5.3 Design of Grids of Beams

5.3.1 Open Questions and Preliminaries

In this section, the open questions and preliminaries are given for the non-adaptive and adaptive algorithms solving the design of grids of beams if reuse medium access is assumed. Since time-frequency units are reuse among APs, the influence of the co-channel interference must be considered. According to Fig. 3.2, the design of grids of beams is processed by the BS. The design of grids of beams is provided by the BS for each reuse group $\mathcal{T}_n$ independently, where $n = 1, 2, ..., N_{\mathrm{SF}}$. These reuse groups are known to the BS. Furthermore, the BS knows the beamforming vectors and the antenna geometry of each AP in the cell such that the BS is aware of the antenna patterns. Patterns with clearly shaped main lobes are assumed, i.e., the magnitude of a main lobe is clearly larger than the one of the side lobes. Only the azimuth is considered.

For the adaptive algorithm, the BS requires an additional knowledge compared to the non-adaptive algorithm. This knowledge is the same one as assumed in Section 4.2.1, where the subproblem is addressed for orthogonal medium access:

- The BS knows the set $\mathcal{R}_{\mathrm{RS}}$ of RSs and the directions from the BS to all RSs of the cell.
- The BS knows instantaneously which UEs are served by an AP t represented by the set $\mathcal{R}_t$.
- The BS knows the PDF $\rho(x, y)$ representing the distribution of the UEs in the cell. For the sake of simplicity, it is assumed that the amount of data sent by an AP is proportional to the number of served UEs.
- The BS knows an appropriate pathloss model [Par00] well suited for the geographic environment in which the relay network operates.

5.3.2 Non-Adaptive Design of Grids of Beams

In this section, the non-adaptive algorithm is presented. The grids of beams applied by the AP t are designed by the non-adaptive algorithm defined in Section 4.2.2. However, one difference exists in order to combat the co-channel interference. The grids of beams are allocated to the time-frequency units randomly and not in a periodical order. The allocation related to an AP is independent from the other APs. If beam b is applied on multiple time-frequency units, it is unlikely due to the random allocation that each time-frequency unit is affected by strong co-channel interference.

5.3.3 Adaptive Design of Grids of Beams

In this section, the adaptive algorithm solving the design of grids of beams is introduced if reuse medium access is considered. The strategy pursued for the reuse medium access is similar to the one for the orthogonal medium access presented in Section 4.2.3: Beams are more often applied if a large number of UEs is expected in the direction of the beams. The grids of beams are created for each reuse group $\mathcal{T}_n$ in two steps. At first, the demand values are determined. Then, beams are grouped in grids of beams and allocated to time-frequency units such that the interference averaged over all positions of receiving stations is mitigated taking into account the demand values.

In the first step, the demand values $D_{t,b}$ are determined for each AP t and each beam b, where $t \in \mathcal{T}_n$ and $b = 1, 2, ...B_t$. The demand value represents for how many time-frequency units a beam is used in a subframe. Since a similar strategy is pursued as chosen for orthogonal medium access, the demand values are determined as introduced in Section 4.2.3: The demand values yield from (4.2) and (4.3) if an AP is an RS or a BS in order to apply beams more often which are expected to be preferred by the UEs and RSs. The constraint that the sum of the demand values is limited according to (4.1) must be fulfilled for each AP.

In the second step, the problem is solved how beams are grouped in grids of beams and how they are allocated to time-frequency units. In order to describe the problem, the definition of the average interference originally given in Section 4.2.3 is extended to the case that co-channel interference exists. Then, the description of the problem as an integer program is given and the solution is presented.

In order to decide if two beamforming vectors applied on the same time-frequency units generate a weak or strong interference, the modelling of the interference and the average interference of an RS and BS is defined. The definition is based on (4.4), (4.5) and (4.6) representing the definitions in orthogonal medium access. The interfering AP is denoted by t'. The beam applied by AP t' is denoted by b'. The interference $I_{(t',b')}(x, y)$ generated by beam b' of AP t' is modelled by (4.4) for position (x, y). Assume that AP t and AP t' are elements of $\mathcal{T}_n$ and AP t is an RS. Then, the interference caused by beam b' of AP t' in the sector of beam b of RS t is modelled by

$$I_{(t',b')}^{(t,b)} = \int\int_{A^{(t,b)}} I_{(t',b')}(x, y)\rho(x, y)\; dxdy, \tag{5.6}$$

where $A^{(t,b)}$ and $\rho(x, y)$ denote the sector served by beam b of RS t and the PDF representing that probability that a UE served by t is at position (x, y). If t and t' represent the same RS, i.e., $t = t'$, equation (5.6) and (4.5) are identical and represent inter-beam interference. If $t \neq t'$, equation (5.6) represents co-channel interference. If the AP t is the BS, the average interference caused by beam b' of AP t' occurres in the coverage area $A^{(0,b)}$ and at the RSs $\mathcal{R}_{\mathrm{RS},b}$ if the main lobe of beam b is directed to these RSs. The average interference is given based on (4.6) in order to consider that the receiving stations of the BS are UEs and RSs. The average interference $I_{(t',b')}^{(0,b)}$ is defined by

$$I_{(t',b')}^{(0,b)} = \frac{|\mathcal{R}_0|}{N_{\mathrm{UE}}} \int\int_{A^{(0,b)}} I_{(t',b')}(x, y)\rho(x, y)\; dxdy + \sum_{r \in \mathcal{R}_{\mathrm{RS},b}} \frac{|\mathcal{R}_r|}{N_{\mathrm{UE}}} I_{(t',b')}(x_r, y_r), \tag{5.7}$$

where $|\mathcal{R}_r|$, N_{UE} and (x_r, y_r) denote the number of receivers of AP r, the number of UEs in the cell and the coordinates of the RS r, respectively.

In total, $\left(\sum\limits_{t\in\mathcal{T}_n} B_t\right)^2$ different average interference values exists for subframe n. As in Section 4.2.3, the average interference values are pooled in a matrix. The matrix is a block diagonal matrix called $\mathbf{A} \in \mathbb{R}^{F\sum_{t\in\mathcal{T}_n} B_t \times F\sum_{t\in\mathcal{T}_n} B_t}$ and contains the average interference for all beams, time-frequency units and APs considered in subframe n. The blocks of the diagonal have a size of $\sum\limits_{t\in\mathcal{T}_n} B_t \times \sum\limits_{t\in\mathcal{T}_n} B_t$ and are filled with the average interference values while the elements outside of the diagonal are zero. Furthermore, the assignment variable $v_{t,b,k}$ is used. All assignment variables are grouped in the vector $\mathbf{v} \in \{0,1\}^{F\sum_{t\in\mathcal{T}_n} B_t \times 1}$. Then, the problem is given by

$$\min_{\mathbf{v}} \mathbf{v}^{\mathrm{T}}\mathbf{A}\mathbf{v} \tag{5.8a}$$

subject to:

$$\sum_{b=1}^{B_t} v_{t,b,f} = G_t, \qquad 1 \le f \le F, t \in \mathcal{T}_n, \tag{5.8b}$$

$$\sum_{f=1}^{F} v_{t,b,f} = D_{t,b}, \qquad 1 \le b \le B_t, t \in \mathcal{T}_n, \tag{5.8c}$$

$$v_{t,b,f} \in \{0,1\}. \tag{5.8d}$$

The integer program has the same structure as (4.7), but its complexity is increased in general since the number of considered APs is not always one, but given by $|\mathcal{T}_n|$. Due to the same structure of the integer programs, the algorithm illustrated in Fig. 4.2 is extended such that all APs of $\mathcal{T}_n$ are considered. Beams are allocated to time-frequency units sequentially, where the best-fitting beam is chosen in each iteration. The extended algorithm is depicted in Fig. 5.1. In the first step, the set $\mathcal{G}_{t,f}$ representing the allocated beams is initialized for each time-frequency unit f and each AP t of reuse group $\mathcal{T}_n$. In the second step, beams are allocated to time-frequency units sequentially. The different APs are considered in the innermost loop in order to prevent that an AP is favored. The beam b^* which has the maximum demand value is chosen. The beam b^* is allocated to the best fitting time-frequency unit f^*. The best fitting time-frequency unit f^* is the time-frequency unit where the beam b^* causes the lowest interference in the sector of the already allocated beams and where the lowest interference occurs in the sector $A^{(t,b^*)}$. The time-frequency unit f^* is searched among all units which are used by AP t and to which not yet G_t beams are assigned. These time-frequency units

---

Step 1: *Initialization*

$\mathcal{G}_{t,f} := \{\}$ for all f and for each $t \in \mathcal{T}_n$

Step 2: *Design grids of beams*

for $f = 1$ to F **do**

  for $g = 1$ to $\max\limits_{t \in \mathcal{T}_n}\{G_t\}$ **do**

    for all $t \in \mathcal{T}_n$ **do**

      if $g \leq G_t$ **do**

        $b^* := \operatorname*{argmax}\limits_{b}\{D_{b,t}\}$

        Choose best fitting time-frequency unit according to (5.9)

        $\mathcal{G}_{t,f^*} := \mathcal{G}_{t,f^*} \cup \{f^*\}$

        $D_{(t,b^*)} := D_{(t,b^*)} - 1$

Figure 5.1. Adaptive design of grids of beams.

are represented by the set $\mathcal{F}_t$. The time-frequency unit f^* is found by

$$f^* = \operatorname*{argmin}_{\mathcal{F}_t} \left(\sum_{t' \in \mathcal{T}_n} \sum_{g \in \mathcal{G}_f} I_{(t,b^*)}^{(t',g)} + I_{(t',g)}^{(t,b^*)} \right). \tag{5.9}$$

In contrast to (4.8), multiple APs of the same reuse groups are considered in (5.9). The beam b^* is assigned to $\mathcal{G}_{t,f^*}$ and its demand value D_{t,b^*} is decreased by one. The sequential allocation is repeated FG_t times for each AP t such that a grid of beams is allocated to each time-frequency unit of each AP of the reuse group.

The computational complexity is given in Table 5.1 if the adaptive algorithm for the design of grids of beams is applied to all subframes of a frame. It is assumed that the PDF $\rho(x, y)$, the number F of time-frequency units, the number G_t of applied beams and the number B_t of beamforming vectors are invariant over the time for all APs, only the number of UEs and their assignment is changed. Then, the computational load is caused by the calculation of the demand values of the BS according to (4.3), the calculation of the average interference values of the BS according to (5.7) and the algorithm depicted in Fig. 5.1. Since the search of time-frequency unit f^* in (5.9) depends on the instantaneous allocation, an upper bound is given for the computational complexity of (5.9) as already provided in Section 4.2.3 for (4.8). It is assumed that the time-frequency unit f^* is always searched among all F time-frequency units and that $G_t - 1$ beams are allready allocated in the set $\mathcal{G}_{t,f}$.

Table 5.1. Computational complexity of the adaptive algorithm for the design of grids of beams.

Multiplications/divisions	$3B_0 + \sum_{b=1}^{B_0} \sum_{t \in \mathcal{T}_1} B_t(\lvert\mathcal{R}_{\mathrm{RS},b}\rvert + 2)$
Additions/subtractions	$\sum_{n=1}^{N_{\mathrm{SF}}} \left(\sum_{t \in \mathcal{T}_n} (FG_t) + (\lvert\mathcal{T}_n\rvert - 1 + \sum_{t \in \mathcal{T}_n} (2G_t - 3))F^2 G_t \right) +$ $\sum_{b=1}^{B_0} \lvert\mathcal{R}_{\mathrm{RS},b}\rvert + \sum_{b=1}^{B_0} \sum_{t \in \mathcal{T}_1} B_t \lvert\mathcal{R}_{\mathrm{RS},b}\rvert$
Search maximum/minimum	for each n: search $G_t F$ times for each $t \in \mathcal{T}_n$ in set of size B_t for each n: search $\sum_{t \in \mathcal{T}_n} FG_t$ times in set of size F

5.4 Initial Allocation of Resource Blocks by the RS

5.4.1 Open Questions and Preliminaries

In this section, the open questions and preliminaries are given for the initial allocation of resource blocks solved by each RS. Each RS transmits in one of the N_{SF} subframes of a frame. For the RS t, the question must be answered which resource blocks of a subframe are allocated to the link (t, r), where $r \in \mathcal{R}_t$. The RS t is aware of

- the subframe size S_n for each subframe n, where $n = 1, 2, ...N_{\mathrm{SF}}$.
- SINR value for each resource block k and for each link (t, r), where $k = 1, 2, ...K_t$ and $r \in \mathcal{R}_t$.
- the SINR threshold $\gamma_{\epsilon,r}$ for each number ϵ of bits, where $\epsilon \in \mathcal{E}$, and for each receiving UE $r \in \mathcal{R}_t$ in order to guarantee the bit error probability $BEP_{t,r}$ for each link (t, r).
- the minimum data rate $R_{\mathrm{min},t,r}$ for each link (t, r) if the objective is to maximize the sum rate as claimed in problem (P2).

Since the power must not be changed by the RS according to Section 3.3, the RS is also aware of the number of bits which can be transmitted by resource block k if k is allocated to link (t, r). The number of bits transmitted per slot and resource block is denoted by $\epsilon_{r,k}$ and is determined by (4.10). The number of bits transmitted in resource block k of subframe n is $\epsilon_{r,k} S_n$. If the allocation of resource blocks to links is determined, each RS t sends an initial data rate value to the BS for each link (t, r), where $r \in \mathcal{R}_t$. The initial data rate value is denoted as $\hat{R}_{t,r}$ for link (t, r).

5.4.2 Non-Adaptive Allocation of Resource Blocks by the RS

The non-adaptive algorithm for the allocation of resource block is the same as introduced in Section 4.3.2. A weighted round robin is applied. The weights are determined according to (4.11) or (4.12) if the objective is the maximization of the minimum user rate and the maximization of the sum rate, respectively.

5.4.3 Adaptive Allocation of Resource Blocks by the RS Aiming at Maximizing the Minimum User Rate

In this section, the adaptive algorithm allocating resource blocks is treated if the objective is the maximization of the minimum user rate according to (P1). The subproblem of the allocation of resource block is rather similar to the one in the distributed concept for orthogonal medium access and described by the integer program in (4.14). In contrast to (4.14), the bits per resource blocks are known since the SINR value of each resource block and the subframe size S_n is given for each subframe $n = 1, 2, ...N_{\mathrm{SF}}$ and the objective function of the integer program is changed by including the subframe size. The constraints are the same as in (4.14): At most one resource block of the same time-frequency unit is allocated to a link and each resource block is allocated exclusively to one link. Having this in mind, the description of the subproblem yields from modifying (4.14) and is represented by the following integer program:

$$\max_{u_{r,k}} \min_{r \in \mathcal{R}_t} w_r \sum_{k=1}^{K_t} u_{r,k} \epsilon_{r,k} S_n \tag{5.10a}$$

subject to:

$$\sum_{k=(f-1)G_t+1}^{fG_t} u_{r,k} \leq 1, \qquad 1 \leq f \leq F, \ r \in \mathcal{R}_t, \tag{5.10b}$$

$$\sum_{r \in \mathcal{R}_t} u_{r,k} = 1, \qquad 1 < k \leq K_t, \tag{5.10c}$$

$$u_{r,k} \in \{0, 1\}, \qquad 1 < k \leq K_t, \ r \in \mathcal{R}_t. \tag{5.10d}$$

Since the variable S_n is not a function of k, r and $u_{k,r}$, it affects which data rate values are achieved, but not which resource blocks are allocated to the links. Hence, the subproblem described by integer program (5.10) can also be solved like (4.14) by the algorithm introduced in Fig. 4.3.

5.4.4 Adaptive Allocation of Resource Blocks by the RS Aiming at Maximizing the Sum Rate

In this section, the subproblem of the allocation of resource blocks is treated with the objective of maximizing the sum rate according to (P2). Since the subframe size is known, the RS t is able to determine the bits per subframe for each link (t, r), where $r \in \mathcal{R}_t$ if the resource blocks are allocated initially. The RS allocated so many resource blocks to each link, that the number of bits transmitted in the subframe are sufficient to achieve the minimum data rate. If the number of bits are not sufficient, a UE must be dropped or the minimum data rate must be reduced according to Section 4.5.3. At first, the subproblem is described as an integer program. Then, the algorithm itself is presented.

The subproblem is similar to the problem of the maximizing the sum rate by allocating resource blocks described by (4.22) and solved by the RSs in orthogonal medium access. The objective is to maximize the sum rate as known from (4.22), but $\epsilon_{r,k}$ is substituted by $\epsilon_{r,k} S_n$ here since the subframe size S_n is known. The objective is restricted to three constraints. Two constraints are also known from (4.22) and represent that at most one resource block of the same time-frequency unit is allocated to a link and that each resource block is allocated exclusively to one link. Additionally, the allocation must ensure that each link (t, r) achieves the minimum data rate $R_{\mathrm{min},r}$. Then, the integer program is defined as

$$\max_{u_{r,k}} \sum_{r \in \mathcal{R}_t} \sum_{k=1}^{K_t} u_{r,k} \epsilon_{r,k} S_n \tag{5.11a}$$

subject to:

$$\sum_{k=(f-1)G_t+1}^{fG_t} u_{r,k} \leq 1, \qquad 1 \leq f \leq F,\ r \in \mathcal{R}_t, \tag{5.11b}$$

$$\sum_{r \in \mathcal{R}_t} u_{r,k} = 1, \qquad 1 \leq k \leq K_t, \tag{5.11c}$$

$$\sum_{k=1}^{K_t} u_{r,k} \epsilon_{r,k} S_n \geq R_{\mathrm{min},r}, \qquad r \in \mathcal{R}_t, \tag{5.11d}$$

$$u_{r,k} \in \{0, 1\}, \qquad 1 < k \leq K_t,\ r \in \mathcal{R}_t. \tag{5.11e}$$

A similar allocation problem is formulated in [ZL04], where the allocation of subcarriers is addressed for the transmission from a BS to multiple UEs. Since spatial multiplexing is not treated in [ZL04], constraint (5.11b) is neglected there. The problem is solved by a greedy algorithm in order to find an allocation with a low computational complexity.

The algorithm consists of multiple iterations. In each iteration, a resource block is allocated to the link identified by the lowest number of bits carried by the resource blocks allocated in previous iterations. The greedy algorithm is design such that the first priority is that each receiver achieves its minimum data rate. If this aim is achieved, remaining subcarriers are allocated with the aim to maximize the sum rate. In this thesis, this strategy is also used to define an adaptive algorithm. A pseudo code of the algorithm is illustrated in Fig. 5.2. The algorithm is organized in four steps. In the first step, the initial data rate value $\hat{R}_{t,r}$ is set to zero for each link (t, r) and the sets $\mathcal{K}$ and $\mathcal{K}_r$ are defined in order to describe the resource blocks available for an allocation and the resource blocks allocated to link (t, r), respectively. Additionally, the set $\mathcal{R}'$ is initialized as a copy of $\mathcal{R}_t$. In the second step, all receivers get their first resource block successively. This step is identical with the second step of the algorithm depicted in Fig. 4.3. In the third step, the resource blocks of set $\mathcal{K}$ are allocated as long as resource blocks remain or until each link achieves its minimum data rate. In each iteration, the receiver r^* with the minimum weighted initial data rate value $w_{r^*}\hat{R}_{t,r^*}$ is selected. The weight w_{r^*} is defined as

$$w_{r^*} = \frac{\hat{R}_{t,r^*}}{\sum\limits_{r\in\mathcal{R}_t} \hat{R}_{t,r}}, \tag{5.12}$$

similar to (4.26). The last step is identical to to the second step of the algorithm shown in Fig. 4.5. The sum rate is increased by allocating resource blocks to the links with the highest SINR values.

The computational complexity required for the adaptive algorithm is given in Table 5.2. The value of the multiplications is given for the worst case assumptions that $w_r \neq 1$ for each receiving station r. If $w_r = 1$, multiplications by the factor w_r can be omitted. The numbers of iterations of Step 2, Step 3 and Step 4 are not constant for a frame since the numbers of iterations depend on the instantaneous number of UEs and CQI values. Hence, it is not fixed how often the receiver r^* with the minimum weighted initial data rate value $w_{r^*}\hat{R}_{t,r^*}$ is selected. Additionally, the number of additions is not fixed. Here, the worst case values are given.

Table 5.2. Computational complexity of the resource block allocation algorithm aiming at maximizing the sum rate.

Multiplications/divisions	$FG_t - \lvert\mathcal{R}_t\rvert$
Additions/subtractions	FG_t
Search maximum/minimum	search FG_t times in set of size FG_t, search $FG_t - \lvert\mathcal{R}_t\rvert$ times in set of size $\lvert\mathcal{R}_t\rvert$

Step 1: *Initialization*

$\hat{R}_{t,r} := 0$ for all $r \in \mathcal{R}_t$
$\mathcal{K} := \{1, 2, ..., K_t\}$
$\mathcal{K}_r := \{\}$ for all $r \in \mathcal{R}_t$
$\mathcal{R}' := \mathcal{R}_t$

Step 2: *Allocate first resource block to each link*

for each $r \in \mathcal{R}_t$ **do**

$k^* := \underset{k \in \mathcal{K}}{\operatorname{argmax}}\{\gamma_{r,k}\}$

$\mathcal{K}_r := \mathcal{K}_r \cup \{k^*\}$

$\hat{R}_{t,r} := \hat{R}_{t,r} + \epsilon_{r,k^*}$

$k^* \mapsto f^*$

$\gamma_{r,k} := -\infty$ for $(f^* - 1)G_t + 1 \leq k \leq f^* G_t$

$\mathcal{K} := \mathcal{K} \setminus \{k^*\}$

Step 3: *Achieve minimum rate*

while $\mathcal{K} \neq \{\}$ **and** $\mathcal{R}' \neq \{\}$ **do**

$r^* := \underset{r \in \mathcal{R}'}{\operatorname{argmin}}\{w_r \hat{R}_{t,r}\}$

$k^* := \underset{k \in \mathcal{K}}{\operatorname{argmax}}\{\gamma_{r^*,k}\}$

$\mathcal{K}_{r^*} := \mathcal{K}_{r^*} \cup \{k^*\}$

$\hat{R}_{t,r^*} := \hat{R}_{t,r^*} + \epsilon_{r^*,k^*}$

$k^* \mapsto f^*$

$\gamma_{r^*,k} := -\infty$ for $(f^* - 1)G_t + 1 \leq k \leq f^* G_t$

$\mathcal{K} := \mathcal{K} \setminus \{k^*\}$

if $\hat{R}_{t,r^*} \geq R_{\min,t,r^*}$ **do** $\mathcal{R}' := \mathcal{R}' \setminus \{r^*\}$

Step 4: *Increase sum rate*

for each $k^* \in \mathcal{K}$ **do**

$r^* := \underset{r \in \mathcal{R}_t}{\operatorname{argmax}}\{\gamma_{r,k}\}$

$\mathcal{K}_{r^*} := \mathcal{K}_{r^*} \cup \{k\}$

$k^* \mapsto f^*$

$\gamma_{r^*,k} := -\infty$ for $(f^* - 1)G_t + 1 \leq k \leq f^* G_t$

Figure 5.2. Resource block allocation algorithm aiming at maximizing the sum rate according to the objective of (P2).

5.5 Allocation of Resource Blocks by the BS

5.5.1 Open Questions and Preliminaries

In this section, the open questions and preliminaries are given for the allocation of resource blocks solved by the BS. The BS transmits in the first subframe of a frame. For the BS $t = 0$, the questions must be answered which resource blocks are allocated to the link $(0, r)$, where $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\text{RS}}\}$. The BS is aware of:

- the subframe size S_n for each subframe n, where $n = 1, 2, ...N_{\text{SF}}$.
- SINR value for each resource block k and for each link $(0, r)$, where $k = 1, 2, ...K_0$ and $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\text{RS}}\}$.
- the SINR threshold $\gamma_{\epsilon,r}$ for each number ϵ of bits, where $\epsilon \in \mathcal{E}$, and for each receiving station $r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\text{RS}}\}$ in order to guarantee the bit error probability $BEP_{0,r}$ for each link $(0, r)$.
- the minimum data rate $R_{\text{min},t,r}$ for each link (t, r) in the cell if the objective is to maximize the sum rate as claimed in problem (P2).
- the initial data rate values $\hat{R}_{t,r}$ for each link (t, r), where $t \in \mathcal{R}_{\text{RS}}$ and $r \in \mathcal{R}_t$.

Using the SINR values and the SINR thresholds in (4.10), the BS determines like an RS the bits which can be transmitted per resource block if a resource block is allocated. These bits are denoted by $\epsilon_{t,k}S_1$. Then, the allocation of resource blocks to the links is performed. Since the values of the subframe sizes and of the initial data rates are known, the BS can allocate the resource blocks with respect to the maximum number of bits which can be forwarded by an RS during the frame.

5.5.2 Non-Adaptive Allocation of Resource Blocks by the BS

In this section, the non-adaptive algorithm applied by the BS is presented. Actually, the weighted round robin algorithm as introduced for the BS in Section 4.4.2 is applied, but it is extended. Based on the subframe sizes and initial data rate values, the BS detects if resource blocks are allocated to an RS although it cannot forward received bits.

Assume that the resource blocks are allocated from $k = 1$ to $k = K_0$ sequentially. If the resource block k' is allocated to the link $(0, r)$, where $r \in \mathcal{R}_{\mathrm{RS}}$, the bits carried by the resource blocks already allocated are summed up and represented by

$$R_{0,r} = \sum_{k=1}^{k'} u_{r,k} \epsilon_{r,k} S_0. \tag{5.13}$$

Assume that the RS r forwads in subframe n, where the subframe size is S_n. The sum of the initial data rate values of the links served by the RS r is given by

$$\hat{R}_r = \sum_{r' \in \mathcal{R}_r} \hat{R}_{r,r'}. \tag{5.14}$$

If so many resource blocks are allocated to RS r that the RS cannot forward all received bits, i.e.,

$$R_{0,r} > \hat{R}_r \cdot S_n, \tag{5.15}$$

the RS t is not considered in the weighted round robin any longer.

5.5.3 Adaptive Allocation of Resource Blocks by the BS Aiming at Maximizing the Minimum User Rate

In this section, the allocation of resource blocks is presented if the objective is the maximization of the minimum user rate according to (P1). The subproblem addressed in this section is rather similar to subproblem of the initial allocation of the resource blocks, which is solved by the RS and described by (5.10). As in the previous Section 5.5.2, differences occur since the BS is aware of the data rate values which can be forwarded by the RSs. This knowledge is included in the formulation of the subproblem derived as follows.

The objective is the maximization of the minimum data rate. This is equivalent to the maximization of the data rate on the link $(0, r')$, where $r' \in \mathcal{R}_0$ and r' is chosen arbitrarily, if two constraints are fulfilled. At first, the data rate of each BS-to-UE link $(0, r)$, where $r \in \mathcal{R}_0$ and $r \neq r'$, is larger or equal to the maximized data rate. Secondly, the data rate on the BS-to-RS link $(0, r)$, where $r \in \mathcal{R}_{\mathrm{RS}}$, shall be a multiple of the maximized one according to the inverse of the weight w_r defined in (4.24) and representing the number of UEs served by the RS. However, the data rate which is forwarded by the RS r in subframe n shall be upper bounded by $\hat{R}_r \cdot S_n$, which represents the bits which can be forwaded by the RS r in a frame, where $\hat{R}_r$ is given by (5.14). Hence, the data rate $R_{0,r}$ shall be larger than or equal to the minimum of the weighted maximized data rate and the upper bound. This is expressed in

$$\min \left\{ \frac{R_{0,r'}}{w_r}; \hat{R}_r \right\} \leq R_{0,r}. \tag{5.16}$$

Additionally, it must be fullfilled that at most one resource block of the same time-frequency is allocated to a link and that each resource block is allocated exclusively. Taking into account that the BS is transmitting in subframe $n = 1$, the subproblem is described by

$$\max_{u_{r,k}} \sum_{k=1}^{K_0} u_{r',k} \epsilon_{r',k} S_1 \tag{5.17a}$$

subject to:

$$\sum_{k=1}^{K_0} u_{r,k} \epsilon_{r,k} S_1 \geq R_{0,r'}, \qquad r \in \mathcal{R}_0, \tag{5.17b}$$

$$\min\left\{\frac{R_{0,r'}}{w_r}; \hat{R}_r\right\} \leq R_{0,r}, \qquad r \in \mathcal{R}_{\mathrm{RS}}, \tag{5.17c}$$

$$\sum_{k=(f-1)G_0+1}^{fG_0} u_{r,k} \leq 1, \qquad 1 \leq f \leq F,\ r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}, \tag{5.17d}$$

$$\sum_{r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}} u_{r,k} = 1, \qquad 1 < k \leq K_0, \tag{5.17e}$$

$$u_{r,k} \in \{0, 1\}, \qquad 1 < k \leq K_t,\ r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}, \tag{5.17f}$$

where the data rate $R_{0,r}$ is a function of $u_{r,k}$ as defined in (2.22). The integer program is rather similar to (5.10) describing the subproblem of the initial allocation of resource blocks by the RS if the objective is the maximization of the minimum user rate. The different objective and constraint (5.17c) occur only since the number of bits, which can be forwarded by an RS, is predefined. Thus, the greedy algorithm depicted in Fig. 4.3 and proposed to solve (5.10) is modified slightly in order to solve the integer program. Following modifications are made:

- The set $\mathcal{R}_t$ is replaced by the union $\{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}$ in order to represent the stations receiving from the BS.
- The link $(0, r)$, where $r \in \mathcal{R}_{\mathrm{RS}}$, is not considered any longer in the third step if $R_{0,r} > \hat{R}_r$ in order to prevent that more bits are transmitted to an RS than it can forward.

5.5.4 Adaptive Allocation of Resource Blocks by the BS Aiming at Maximizing the Sum Rate

In this section, the allocation of resource blocks is considered with respect to the objective of maximizing the sum rate given by (P2). It is assumed that each BS-to-UE link and each BS-to-RS link achieve its minimum data rate. Otherwise, a UE must be dropped or the minimum data rate must be reduced according to Section 4.5.3. The subproblem is described by modifying integer program (5.11) representing the initial allocation of resource blocks by the RS if the objective is the maximization of the sum rate. Since the data-rate on the RS-to-UE links is known, the constraint is introduced that so many resource blocks are not allocated to an RS that it receives more bits than it can forward. Then, integer program (5.11) is modified to

$$\max_{u_{r,k}} \sum_{r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}} \sum_{k=1}^{K_0} u_{r,k} \epsilon_{r,k} S_0 \tag{5.18a}$$

subject to:

$$\sum_{k=(f-1)G_0+1}^{fG_0} u_{r,k} \leq 1, \qquad 1 \leq f \leq F,\ r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}, \tag{5.18b}$$

$$\sum_{r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}} u_{r,k} = 1, \qquad 1 \leq k \leq K_0, \tag{5.18c}$$

$$\sum_{k=1}^{K_0} u_{r,k} \epsilon_{r,k} S_0 \geq R_{\min,0,r}, \qquad r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}, \tag{5.18d}$$

$$\sum_{k=1}^{K_0} u_{r,k} \epsilon_{r,k} S_0 \leq \hat{R}_r, \qquad r \in \mathcal{R}_{\mathrm{RS}}, \tag{5.18e}$$

$$u_{r,k} \in \{0, 1\}, \qquad 1 < k \leq K_0,\ r \in \{\mathcal{R}_0 \cup \mathcal{R}_{\mathrm{RS}}\}. \tag{5.18f}$$

The objective function representing the maximization of the bits transmitted by the BS includes the subframe size. Constrains (5.18b) and (5.18c) represent that at most one resource block of the same time-frequency unit is allocated to a link and that each resource block is allocated exclusively. Constraint (5.18d) claims that each link achieves its minimum data rate. Constraint (5.18e) ensures that an RS does not receive more than it can forward. Since the integer program is rather similar to (5.11), the subproblem is solved by nearly the same greedy algorithm as depicted in Fig. 5.2. The only difference is that the BS-to-RS link is only considered in the last step of the algorithm as long as the RS keeps able to forward all received bits.

5.6 Final Allocation of Resource Blocks and Bits

When the allocation of resource blocks is finished by the BS, each RS receives as much bits as it can forward or less. Due to the definition of the algorithms in Section 5.4 and Section 5.5 an RS will not receive more. If the RS receives less bits than it can forward, the data rates of the RS-to-UE links is reduced by selecting more robust modulation and coding schemes and by applying zero padding according to Section 4.6.

5.7 Summary

In this chapter, the subproblems of the allocation of slots, the design of grids of beams and the allocation of resource blocks are introduced for the distributed concept for reuse medium access. For all subproblems, non-adaptive algorithms are presented. Novel algorithms enabling an adaptive allocation at a low complexity are proposed except for the allocation of slots. The adaptive algorithms are related to the objectives defined in (P1) and (P2). As it is also the case in the distributed concept for orthogonal medium access, the adaptive algorithm can be replaced by the corresponding non-adaptive algorithms. However, this is paid by an expected performance loss.

Chapter 6
Performance Evaluation

6.1 Introduction

In this Chapter, the proposed concepts are evaluated. At first, the signalling overhead required by the concepts is analyzed by a comparison to a central solution in which all subproblems are solved by the BS. Then, the performance of the adaptive algorithms related to the design of grids of beams, allocation of resource blocks, allocation of power and bits and allocation of slots are analyzed by a comparison to the non-adaptive ones. The performance measure of the comparison is the average minimum user rate in the cell in order to reflect the objective of maximizing the minimum user rate as considered in (P1). The average sum rate is measured and it is also measured how often all UEs served in a frame achieve their minimum user rate if the analysis reflects the objective of (P2) given as the maximization of the sum rate subject to minimum user rate values. Since the adaptive and non-adaptive algorithms differ in the required computational complexity, values of the computational complexity are presented in order to complete a fair comparison.

The chapter is organized as follows: The distributed concept for orthogonal medium access and the distributed concept for reuse medium access are evaluated in Section 6.2 and Section 6.3, respectively. Both sections are divided further and have the same structure. At first the scenario considered for the evaluation is introduced. Then, the entire concept is analyzed concerning signalling overhead, average minimum user rate and average sum rate. In the following, the algorithms related to the design of grids of beams, allocation of resource blocks, allocation of power and bits and allocation of slots are analyzed in the same order as introduced in Chapter 4 and Chapter 5. Finally, the summary of the the main results of the performance evaluation is given.

6.2 Distributed Concept for Orthogonal Medium Access

6.2.1 Evaluation Scenario

The distributed concept for orthogonal medium access is evaluated by snapshot simulations. In a snapshot, random positions of the UEs and random channel transfer functions for the links of a cell are generated in a predefined scenario. One exemplary scenario is considered in order to give a proof of concept and to evaluate the impact of the adaptive algorithms related to the non-adaptive ones. In a snapshot, one frame is considered. For the next snapshot, new positions of the UEs and new channel transfer functions are generated in the exemplary scenario. The definition of the exemplary scenario is given in this section.

The deployment of the APs is illustrated in Fig. 6.1. The cell has the size of three neighbored hexagons. There are one BS and two RSs. The deployment is kept rather simple, but it covers various possible applications of RSs as introduced in Section 1.1. For instance, the deployment is covered in which

- a small number of RSs are deployed since RSs are used to transmit in coverage holes,
- a large number of RSs are deployed but only two RSs are transmitting within the frame considered in the snapshot.

In each snapshot, the UEs are assigned to the BS or an RS by a simple method taken from [MKWK07a] and explained in the following. Three assumptions are made for the assignment of the UEs. These assumptions are only valid for the assignment in order to keep the method rather simple: The power is assumed to be allocated uniformly among all resource blocks. An omnidirectional antenna pattern is assumed since the antenna gain as a result of the design of grids of beams and allocation of resource blocks is not known yet. Spatial multiplexing is neglected for simplicity. Based on these assumptions, the Signal-to-Noise Ratio (SNR) value averaged over all resource blocks is determined for each link. The SNR value is a random variable since a statistical model is applied for the channel modelling. Shannon's channel capacity formula [Bla84] and its extensions to two-hop connections [Doh04] are applied to find the BS or the RS promising the highest user rate. The UE is assigned to the BS or RS promising the highest data rate. In the exemplary scenario, UEs are randomly placed such that each postion is equally probable. The dotted area depicted in Fig. 6.1 represents the area in which a UE is most likely assigned to the BS. In the other area, a UE is most likely assigned to the RS given by the shortest distance.

The scenario is further defined by the parameters listed in Table 6.1. The parameters are not conforming to a standard, but specify a general OFDMA-based network representing basic features of a network according to IEEE 802.16 [IEE04] or LTE [3GP06]. These parameters are valid for the whole Section 6.2 if not stated differently. A bandwidth as supported in [IEE04, 3GP06] is chosen. The number N_{SC} of subcarries in a time-frequency unit and the number F of time-frequency units in a subframe are

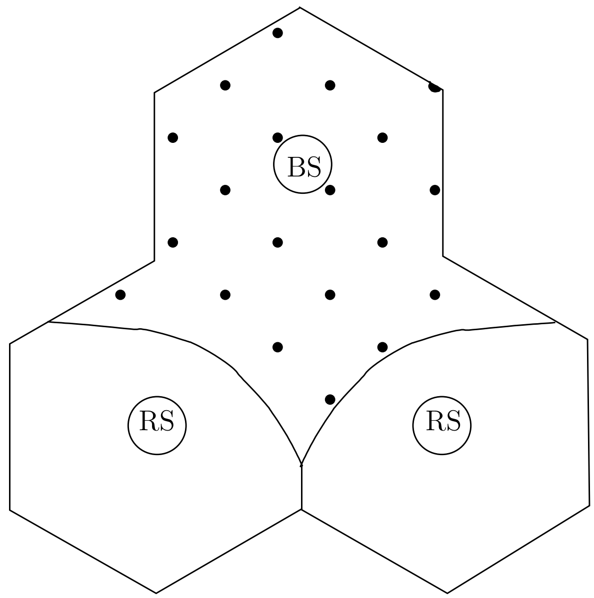

Figure 6.1. Deployment and assignment in the evaluation scenario.

chosen with respect to the coherence bandwidth of the channel as claimed in Section 2.3. A delay spread of the channel which is less than $5 \cdot 10^{-6}$ s is assumed. The number S of slots is chosen with respect to the coherence time as claimed in Section 2.3 if pedestrian users and a carrier frequency of 5 GHz are assumed. The power P_t of the AP t and the noise power given in Table 6.1 are chosen according to [BHIT05]. The set $\mathcal{E}$ of bits per OFDM symbol is chosen such that a wide SINR range is covered. The SINR γ_ϵ required to transmit ϵ bits is taken from Shannon's capacity formula [Bla84] and given by

$$\gamma_\epsilon, r = 2^\epsilon - 1 \tag{6.1}$$

for each receiving station r. Note that these values are chosen for simplicity. A perfect code in terms of the channel capacity and a bit error probability of zero is assumed. In order to analyze the performance in a system with a non-perfect code and a bit error probability larger than zero, the set $\mathcal{E}$ and the corresponding SINR values and bit error probability values are found by link level analysis as given for IEEE 802.16 and LTE in [Hoy04] and [SEW07], respectively. The APs are equipped with a UCA. The UCA consists of 12 dipoles. Only the azimuth is considered. As already assumed in Section 4.2.2, the beamforming vectors used to design the beams are chosen according to [LL00]. The pattern of the beam corresponding to the main lobe direction of 180° is illustrated in Fig. 4.1. The beams corresponding to the main lobe directions 240° and 300° are directed to the RSs. The channel models include a model for fast fading, slow fading and pathloss. The models have been developed within the research project EU IST-4-027756 WINNER II [IST] and are documented in [IST07b]. The BS-to-RS links

Table 6.1. Parameter setting in the evaluation scenario for the orthogonal medium access.

Parameter	Value
bandwidth	5 MHz
number N_{SC} of subcarriers in a time-frequency unit	4
number F of time-frequency units per subframe	64
number S of slots per frame	100
power P_t of BS and RS	35 dBm
noise power	-102 dBm
set $\mathcal{E}$ of bits per OFDM symbol	$\{0, 0.5, 1, 2, 3, 4, 5, 6, 7, 8\}$
antennas at BS and RS	UCA, 12 elements
beam type	Chebyshev, 20 dB side lobe attenuation
main lobe direction	0°, 30°, 60°, ..., 330°
channel model BS to UE and RS to UE	WINNER Chanel Model C2 NLOS [IST07b]
channel model BS to RS	WINNER Chanel Model B5a [IST07b]
outer radius of hexagons	150 m
default value for the number N_{UE} of UEs	20
default value for the number G_t of beams in the grids of beams	3
default value for the minimum user rate $R_{\mathrm{min},0,r}$	10 bits/slot

are modeled as LOS links between two fixed stations deployed over the rooftops in a typical urban area. The other links are modeled as Non-Line-Of-Sight (NLOS) links from a fixed station to mobile UEs in a typical urban area. The number N_{UE} of UEs, the number G_t of beams in the grids of beams and the minimum user rate $R_{\mathrm{min},0,r}$ are varied in the following sections. If the impact of one of these parameters is evaluated, the other two parameters are chosen according to the default values given in the table if not stated differently. Note that the minimum user rate is only considered if the analysis reflects the objective of (P2) given as the maximization of the the sum rate subject to minimum user rate values.

In this thesis, the data rate, user rate and sum rate values are given in bits per slot. For instance, the data rate of the link (t, r) is measured by counting all the bits transmitted by AP t to receiving station r during a frame and by dividing by the number S of slots in a frame. Note that the representation in bits per slot allows a more general representation than in bits per second since the guard intervals introduced by the OFDM transmitters and removed by the OFDM receivers can be neglected.

6.2.2 Concept Evaluation

6.2.2.1 Maximizing the Minimum User Rate

In this section, the distributed concept for orthogonal medium access is evaluated if the objective is the maximization of the minimum user rate according to problem (P1). Firstly, the concept is evaluated concerning the signalling overhead. Then, the concept is evaluated concerning the average minimum user rate chosen with respect to the objective. Finally, the computational complexity is given for the concept.

In Section 3.2, the decomposition of problem (P1) according to the distributed concept for orthogonal medium access is motivated such that an optimum solution of (P1) requires a central solution based on a global knowledge of CSI and noise power values. This central solution requires a large signalling overhead from the RSs to the BS. In the following, this motivation is approved by a comparison of the signalling overhead required for a central solution and for the proposed concept. The signalling overhead sent from the RSs to the BS is derived for the central solution and for the proposed concept.

The signalling overhead required for a central solution is derived as follows. Let us assume that the BS is the central station finding the solution and all RSs signal the CSI and the noise power to the BS since the BS is the only AP directly linked to all RSs. In order to obtain global knowledge, one value representing the noise power and two values representing the real and imaginary part of the complex valued CSI must be transmitted to the BS for each RS-to-UE link in the cell and for each time-frequency unit. Additionally, it must be considered that one value representing the channel gain value of the link between a receiving station and an interfering AP must be sent to the BS for each time-frequency unit. For the RS t, the number N_{Val} of values to be fed back to the BS is

$$N_{\mathrm{Val}} = F|\mathcal{R}_t| + 2F|\mathcal{R}_t| + F|\mathcal{R}_t|(|\mathcal{T}_n| - 1), \tag{6.2}$$

where F, $\mathcal{R}_t$ and $\mathcal{T}_n$ denote the number of time-frequency units, the set of served UEs and the reuse group transmitting in subframe n. The three summands of (6.2) correspond to the noise values, the CSI of the RS-to-UE links and the channel gain values of the $(|\mathcal{T}_n| - 1)$ interferers in the reuse group $\mathcal{T}_t$, respectively. Since orthogonal medium access is considered in this section, the last summand is zero and (6.2) simplifies to

$$N_{\mathrm{Val}} = 3F|\mathcal{R}_t|. \tag{6.3}$$

These values are transmitted to the BS once per frame. This assumption is in line with the assumption made in Section 2.3 that the duration of a frame is smaller than the coherence time. Additionally, this assumption allows the smallest possible coherence time being equal to the duration of a single frame. Since these values represent continuous quantities, these values must be quantized before they are transmitted. A quantization is assumed with Q bits. The number of bits required to be transmitted from RS t to the BS once per frame is

$$N_{\mathrm{MAX}} = QN_{\mathrm{Val}} = 3QF|\mathcal{R}_t|. \tag{6.4}$$

The signalling overhead required for the distributed concept for orthogonal medium access is derived as follows. The RS t sends the initial data rate value for each link (t, r), where $r \in \mathcal{R}_t$. The initial data rate of a link is given per slot and depends on the number F of time-frequency units in a subframe, the number N_{SC} of subcarriers in a time-frequency unit and the maximum number $\max\{\mathcal{E}\}$ of bits carried by an OFDM symbol. The number of bits required to represent the initial data rate values of all $r \in \mathcal{R}_t$ is upper bounded by

$$N_{\text{DC-OMA}} = |\mathcal{R}_t| \cdot \log_2(F \cdot N_{\mathrm{SC}} \cdot \max\{\mathcal{E}\}). \tag{6.5}$$

These bits have to be transmitted from RS t to the BS once per frame.

In order to compare the distributed concept for orthogonal medium access and the central solution, the quantization $Q = 4$ is assumed, where this assumption is rather optimistic for the central solution since only 16 quantization steps are possible in order to represent continuous CSI and noise power values. Fig. 6.2 and Fig. 6.3 show the signalling overhead in bits per slot sent from one RS to the BS required for a central solution found by the BS and for the distributed concept for orthogonal medium access denoted by “DC-OMA“. According to the parameters listed in Table 6.1, 2048 bits per slot can be transmitted on a link at maximum, i.e., if each subcarrier is allocated $\max\{\mathcal{E}\} = 8$ bits per slot. This value acts as a reference. Fig. 6.2 shows the number of signalled bits per slot as a function of the number of UEs in the set $\mathcal{R}_t$. For the central solution and for the distributed concept for orthogonal medium access, the number of bits per slot are proportional to the number of UEs according to (6.3) and (6.5). The central solution requires two decades more bits per slots since CSI is transmitted for each UE and each time-frequency unit. Fig. 6.3 shows the number of signalled bits per slot as a function of the number of time-frequency units per subframe. The number of UEs served by the RS is five and ten, respectively. Fig. 6.3 illustrates that the number of bits per slot are proportional to the number of time-frequency units F for the central solution but only proportional to $\log_2(F)$ for the proposed concept also according to (6.3) and (6.5). In both figures, the signalling overhead required for the central solution is only one or two decades smaller than the reference. This shows that roughly 1% -10% of the time-frequency units being available for the downlink are

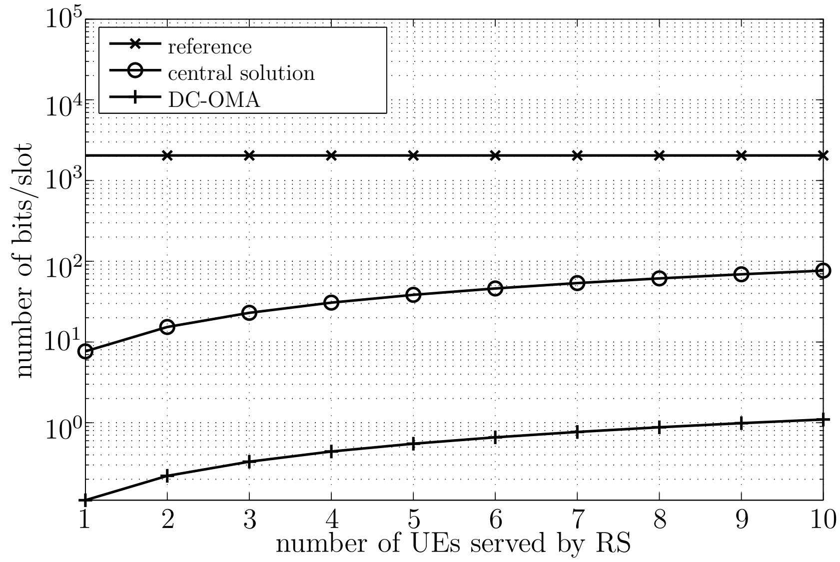

Figure 6.2. Signalling from RS to BS depending on the number of UEs served by the RS.

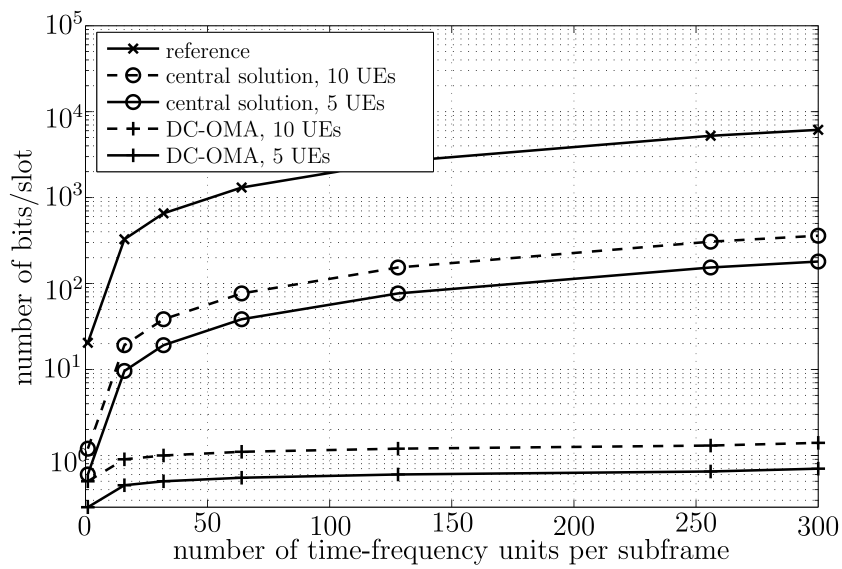

Figure 6.3. Signalling from RS to BS depending on the number of time-frequency units per subframe.

required in order to transmit the signalling overhead corresponding to a single RS if the best performing modulation and coding scheme can be applied for the feedback. If multiple RSs are considered, the signalling overhead increases further as a linear function of the number of RSs. Hence, a central solution is impractical while the proposed concept is rather efficient in terms of signalling overhead.

In the following part of this section, the distributed concept for orthogonal medium access is evaluated concerning the average minimum user rate. Since each subproblem can be solved by an adaptive or a non-adaptive algorithm, various combinations of adaptive and non-adaptive algorithms exist, where each combination may serve as an implementation of the concept. A specific combination is called instance. In this section, two different instances are considered. The first one in which each subproblem is solved by an adaptive algorithm is called all-adaptive instance. The second instance in which each subproblem is solved by a non-adaptive algorithm is called non-adaptive instance. The all-adaptive instance is compared to the non-adaptive instance serving as a lower bound. Related to the performance, an optimum instance is a combination of algorithms solving optimally all integer programs describing all subproblems formulated in Section 4.2 to 4.5. However, such an instance is not applicable due to the intractable computational complexity as stated in the corresponding sections. Hence, an optimum instance is not considered. The performance is evaluated depending on the number of beams applied by the BS and RSs in the grids of beams and the number of UEs. The number of beams affects the interference and the number of resource blocks in the cell. The number of UEs has an impact on the competition for the available resources among the UEs.

In Fig. 6.4, the average minimum user rate as a function of the number of beams used by the BS is shown. For the all-adaptive instance (all-adapt.), the number of beams used by the RSs in a grid of beams are one, two, three, four and six. Additionally, the average minimum user rate is given for the non-adaptive instance (non-adapt.). For the non-adaptive instance, the number of beams used by the RSs in a grid of beams is always six beams since it is the best performing non-adaptive instance. The results of the all-adaptive instance reveal that the average minimum user rate increases for the considered number of beams if more beams are used at the BS and at the RSs since more resource blocks are available. However, the slope decreases if more beams are applied. More beams lead to stronger inter-beam interference since the inter-beam interference caused by the pre-defined beamforming vectors is accumulated. Using all 12 beams in the grids of beams is not considered since the inter-beam interference prevents a successful transmission of bits. The question arises what is the optimum number of beams in the grids of beams. This question must be answered for each particular scenario individually. For instance, increasing the number of beams used by the BS and RS from three to six provides a gain of approximately 40% in terms average minimum user rate. But it is doubtful to use 6 instead of 3 beams because of three reasons. At first, it must be considered that more CQI values must be estimated if the number of beams increases. This requires a larger computational complexity at the receiving UEs and RSs. Secondly, the pilot phase introduced in Section 3.3 must be extended since the CQI values of more resource blocks must be determined by the receiving stations. Thirdly, more CQI values must be fed from the receiving stations to the transmitting APs. Hence, a larger feedback phase also introduced in Section 3.3 is required. However, a larger pilot and feedback phase reduces the time in which the bandwidth is exploited for the actual transmission of frames. These drawbacks are not considered in the results depicted in Fig. 6.4.

Additionally, Fig. 6.4 gives the comparison of the all-adaptive instance to the non-adaptive one. The performance of the non-adaptive instance shows that the user rate of the weakest UE is close to zero. The non-adaptive instance does not gain from the increased number of beams since most of the resource blocks allocated to a UE are based on beamforming vectors not steering in the direction of the UE. Since resource blocks are allocated in a round robin manner and without taking into account CQI values, only one out of 12 resource blocks is based on a beamforming vector generating a beam steered to the UE. These results lead to the conclusion that a significant gain in terms of minimum user rate is achieved if adaptive algorithms are applied.

In Fig. 6.5, the average minimum user rate is illustrated depending on the number of UEs for the all-adaptive instance (all-adapt.) and for the non-adaptive instance (non-adapt.). The same number of beams is used by the BS and RSs. The combinations where two, three and four beams are used in the grids of beams are considered. Since the previous results show that the non-adaptive instance does not benefit from a large number of beams, only a single value is considered for the number of beams. As an example, the default value of three beams is chosen for the non-adaptive instance. The average minimum user rate is inversely proportional to the number of UEs since less resource blocks are available per UE if the number of UEs increases. Fig. 6.5 confirms the result that the gain in terms of average minimum user rate reduces with increasing number of beams. The average minimum user rate of the non-adaptive concept converges quite fast to zero.

In total, the results depicted in Fig. 6.4 and Fig. 6.5 approve that the all-adaptive instance achieves large gains compared to the non-adaptive one, but these gains are paid by additional computational complexity. While the computational complexity

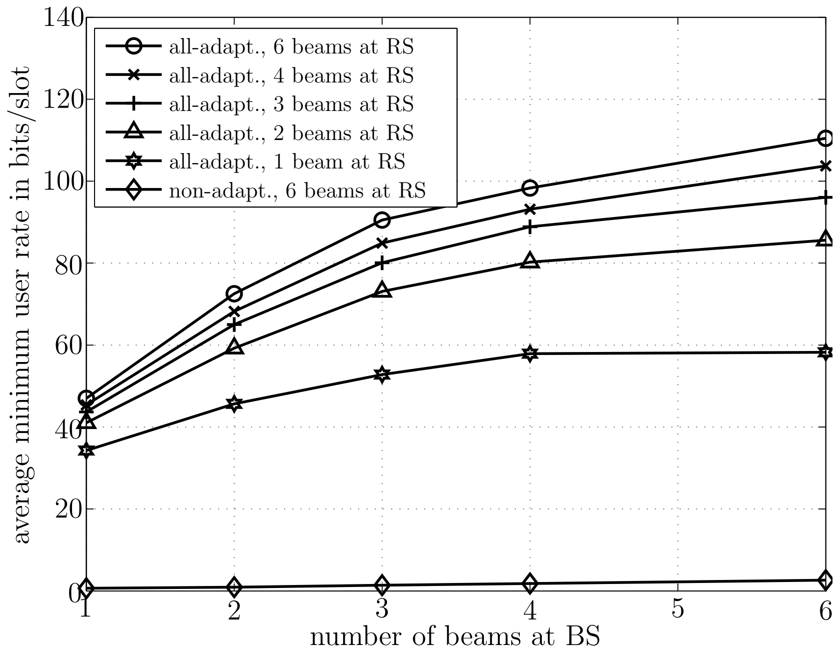

Figure 6.4. Average minimum user rate of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.) as a function of the number of beams used by the BS, various numbers of beams at the RSs for the all-adaptive instance, 6 beams at the RSs for the non-adaptive instance, $N_{\mathrm{UE}} = 20$.

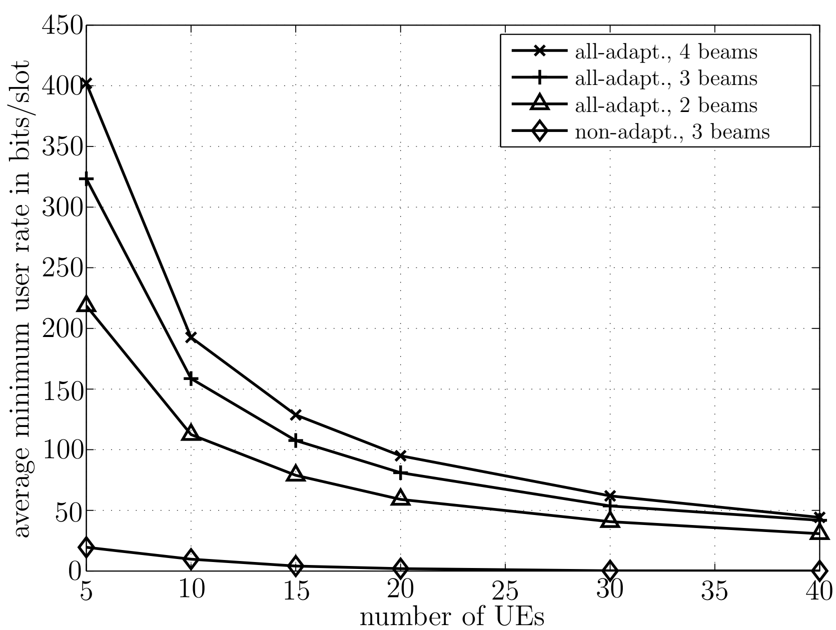

Figure 6.5. Average minimum user rate of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.), same number of beams at the BS and RSs, various numbers of beams for the all-adaptive instance, 3 beams for the non-adaptive instance.

is insignificant for the non-adaptive instance, the all-adaptive instance requires the computational complexity given for each adaptive algorithm applied in the all-adaptive instance in Table 4.2 to Table 4.4 as the number of required multiplications, additions and searches of a maximum or minimum. The computational complexity required for the allocation of slots is derived from (4.45) and (4.46). Again, the exemplary scenario introduced in Section 6.2.1 is considered. The number G_t of beams is given by the default value, i.e., $G_t = 3$ for all t. For the BS, the computational complexity is given in Table 6.2. The assumption is made that ten UEs are assigned to the BS. For an RS, the computational complexity is given in Table 6.3, where five UEs are assumed to be assigned to the RS. Two values are given for the searches of the maximum or minimum. The first value represents the number of repetitions, the second value denotes the size of the set in which the extremum is searched. The computational complexity given in both tables must be provided once per frame.

Table 6.2. Computational complexity per frame of the all-adaptive instance for the BS, where $|\mathcal{R}_0| = 10$.

Adaptive algorithm	Multiplications/ divisions	Additions/ subtractions	Search maximum/ minimum (repetitions, size of set)
Design of grids of beams	348	111194	$(576, 12)$ $(576, 64)$
Allocation of resource blocks	191	192	$(192, 192)$; $(180, 12)$
Allocation of power and bits	84671	39744	$(1728, 12)$; $(1728, 64)$
Allocation of slots	26	9	$(0, 0)$

Table 6.3. Computational complexity per frame of the all-adaptive instance for RS t, where $|\mathcal{R}_t| = 5$.

Adaptive algorithm	Multiplications/ divisions	Additions/ subtractions	Search maximum/ minimum (repetitions, size of set)
Allocation of resource blocks	191	192	$(192, 192)$; $(187, 5)$
Allocation of power and bits	84671	39744	$(1728, 5)$; $(1728, 64)$

The computational complexity is related to today's processor by an example. Assume that a frame has a size of $6 \cdot 10^{-3}$ s and that a multiplication, division, addition or subtraction corresponds to a single floating point operation. Assume further that each search of an extremum in a field of length N requires N floating points operations. In total, the all-adaptive instance requires approximately 0.064 GigaFLOPS (10^9 Floating Point Operations Per Second) processed by the BS and approximately 0.047 GigaFLOPS processed by an RS. For comparison, the processor of a today's personal computer is able to process multiple GigaFLOPS [CKL+08].

In both tables, the comparison between the algorithms show that adaptive algorithms for the allocation of power and bits and for the design of grids of beams require a large computational complexity. The computational complexity of the allocation of resource blocks is rather low. The computational complexity required for the allocation of slots is close to zero since analytical expressions are derived for the sizes of the subframes.

6.2.2.2 Maximizing the Sum Rate

In this section, the evaluation of the distributed concept for orthogonal medium access is presented if the objective is the maximization of the sum rate according to (P2). At first, the signalling overhead is analyzed. Then, the performance of the all-adaptive instance is compared to the non-adaptive one and the computational complexity is shown.

The signalling overhead is analyzed in the following. Like Section 6.2.2, the concept is compared to the central solution in which the BS is the central station finding the optimum solution of (P2) based on full knowledge of CSI and noise power values. The number of bits required to transmit the CSI and noise power values from one RS to the BS is given by (6.4) for the central solution. Hence, the results presented for the central solution in Fig. 6.2 and Fig. 6.3 are also valid here. If the maximization of the sum rate is addressed, the number of bits required for the distributed concept for orthogonal medium access is twice the number given by (6.5) because two subframes are allocated to each RS. Hence, the results presented in Fig. 6.2 and Fig. 6.3 and related to the distributed concept for orthogonal medium access must be doubled if the objective is related to (P2). Despite the factor of two, the same conclusion can be drawn as in the previous Section. The distributed concept for orthogonal medium access reduces the signalling overhead by approximately two decades compared to the central solution.

In the following part of this section, the concept is evaluated by a comparison of the all-adaptive instance and the non-adaptive one. Whenever the objective is the maximization of the sum rate, the performance of an instance is measured by two quantities in this thesis. The first one is the average sum rate. According to the definition of (3.5) and its constraints, (P2) is only solved if all links in the cell achieve their minimum data rate. The average sum rate considers only snapshots in which all links achieve their minimum data rate. However, the performance of the concept is only described completely if the remaining snapshots are considered additionally. Hence, a second quantity called relative frequency of outages is introduced. An outage is defined as a snapshot in which the minimum user rate is not achieved by all links in a cell. The relative frequency of outages is the ratio of the number of outages and the total number of snapshots in a simulation. The performance of an instance is described completely by the value of the average sum rate and of the relative frequency of outages. The non-adaptive instance is an exception. As one could expect from the results of Section 6.2, the non-adaptive instance does not provide the minimum user rate at least to one UE in a snapshot. In order to allow a comparison to other instances, the average sum rate of the non-adaptive instance is the sum rate averaged over all simulated snapshots, while the minimum data rate constraint is neglected.

In the following, the average sum rate is analyzed depending on the number of beams applied by the BS and RSs in the grids of beams, the number of UEs and the minimum user rate. The relative frequency of outages is analyzed depending on the minimum user rates.

In Fig. 6.6, the average sum rate is illustrated as a function of the number of beams applied in the grids of beams. For the non-adaptive instance, the RSs apply always six beams. Except for the non-adaptive instance, the relative frequency of outages is always close to zero. Hence, the average sum rate is not affected by outages. The average sum rate achieved by the all-adaptive instance behaves similar to the average minimum user rate depicted in Fig. 6.4. The average sum rate increases if the BS applies more beams. However, the gain is rather small if the RS applies more beams while the number of beams applied by the BS is kept unchanged. The reason is that the sum rate is mainly affected by the UEs having established a direct connection since these UEs do not require two transmission as the UEs having established two-hop connections. The non-adaptive instance cannot gain from the increased number of beams. Therefore, a huge gain is revealed for the all-adaptive instance compared to the non-adaptive one especially if the number of beams is rather large.

In Fig. 6.7, the average sum rate depending on the number of UEs is given for the all-adaptive instance and the non-adaptive one. All APs apply two, three or four beams in the grids of beams except for the non-adaptive instance, where the APs apply three beams. The relative frequency of outages is close to zero except for the non-adaptive instance. The characteristic of the average sum rate is explained for the case where two beams are applied in the grids of beams. For a small number of

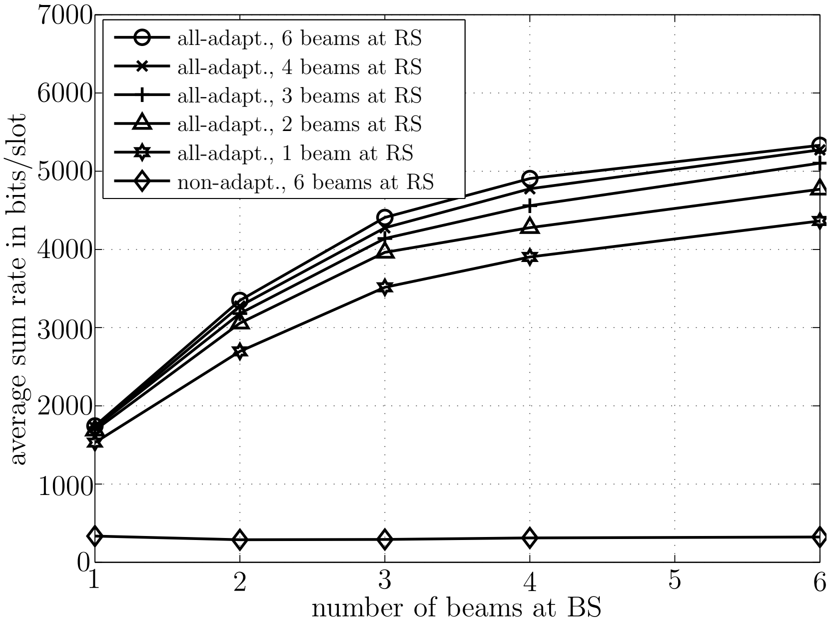

Figure 6.6. Average sum rate of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.) as a function of the number of beams used by the BS, various numbers of beams at the RSs for the all-adaptive instance, 6 beams at the RSs for the non-adaptive instance, $N_{\mathrm{UE}} = 20$, $R_{\min} = 10$ bits/slot.

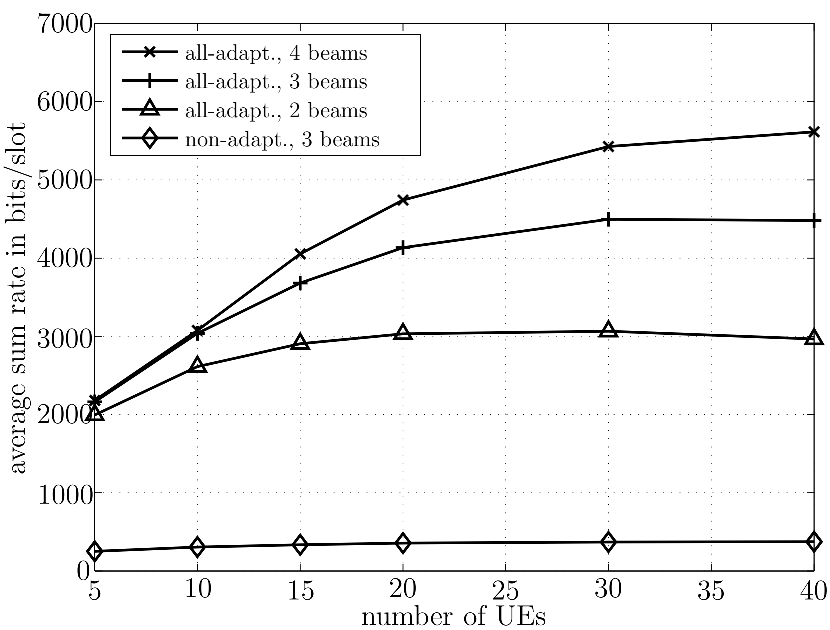

Figure 6.7. Average sum rate of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.), same number of beams at the BS and RSs, various numbers of beams for the all-adaptive instance, 3 beams for the non-adaptive instance, $R_{\min} = 10$ bits/slot.

UEs in the cell, the average sum rate increases since more UEs are available at the same time and the chance increases to allocate a resource block to a UE having a good CQI value. This effect is well known as multiuser diversity [LL05]. However, the average sum rate decreases if more than 30 UEs are considered. Since the minimum user rate must be achieved for each UE, less resource blocks and power can be allcoated to the UE just promising the highest user rate. Except for the non-adaptive instance, this characteristic of the average sum rate depending on the number of UEs is always valid if the objective is the maximization of the sum rate. If the APs apply three or four beams, the characteristic is not shown completely in Fig. 6.7, i.e., the interval in which the average sum rate decreases misses since larger number of UEs are not considered in order to ensure a relative frequency of outages close to zero.

As one can expect from the results of Fig. 6.7, the minimum user rate has a strong impact on the average sum rate. Fig. 6.8 reveals the impact of the minimum user rate. As a reference, the results of the non-adaptive instance are given. Each AP applies three beams in the grids of beams. The average sum rate decreases for the all-adaptive instance if the minimum user rate increases. A minimum user rate larger than 50 bits/slot is not considered as the relative frequency of outages becomes too strong. A minimum user rate of more than 50 bits/slot leads to a relative frequency of outages of more than 5%. Then, the illustration of the average sum rate becomes too optimistic since the average sum rate does not decrease further as only snapshots in which UEs are positioned closely to the APs affect the average sum rate. The non-adaptive instance is not affected by the minimum user rate since it fails nearly every time.

In Fig. 6.9, the relative frequency of outages depending on the minimum user rate is shown for the same setup. As mentioned above, the relative frequency of outages increases for the all-adaptive instance if the minimum user rate exceeds 50 bits/slot. The non-adaptive instance cannot offer a low relative frequency of outages.

As already stated in Section 6.2.2, the gains of the all-adaptive instance compared to the non-adaptive one are paid by additional computational complexity. The computational complexity of the non-adaptive instance is insignificant. The computational complexity required for each adaptive algorithm applied in the all-adaptive instance is given in Table 4.2 to Table 4.9 as the number of required multiplications, additions and searches of a maximum or minimum. If the all-adaptive instance is applied, the computational complexity required for the BS and the RS is given in Table 6.4 and Table 6.5, respectively. The number G_t of beams is assumed to be the default value, i.e., $G_t = 3$ for all t. The assumption is made that ten

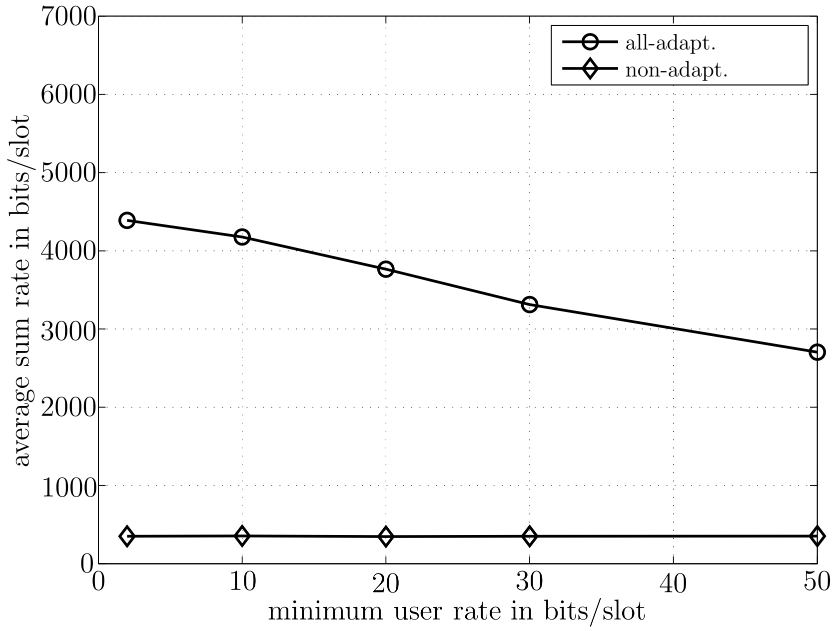

Figure 6.8. Average sum rate of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.), $G_t = 3$ for all APs, $N_{\mathrm{UE}} = 20$.

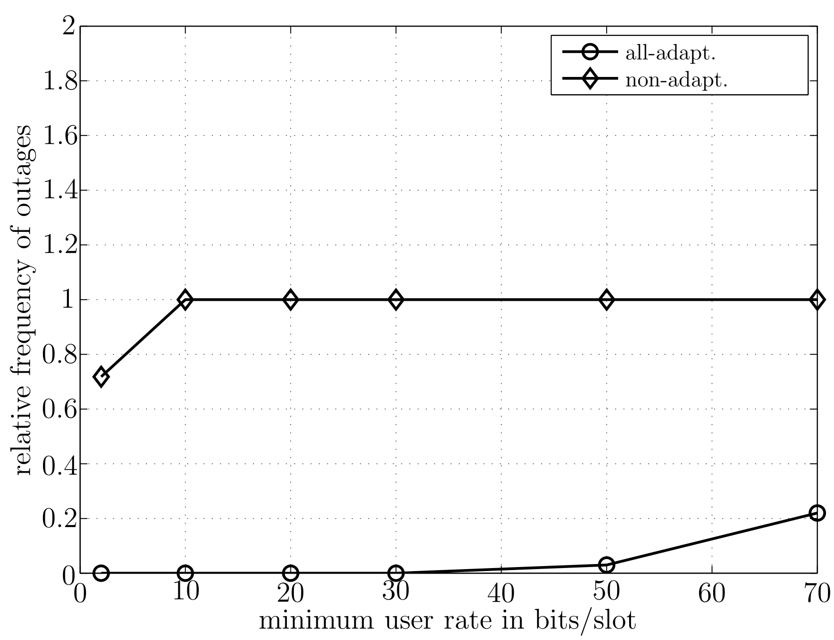

Figure 6.9. Relative frequency of outages of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.), $G_t = 3$ for all APs, $N_{\mathrm{UE}} = 20$

UEs are assigned to the BS and five UEs to a RS. The computational complexity for the allocation of slots depends on the number $S_{\max}$ of slots being available for the maximization of the sum rate. Here, the worst case assumption is made that $S_{\max} = S - 1 = 99$ slots. For each AP a first and a second allocation of resource blocks, power and bits is considered. The first one aims at maximizing the minimum user rate, the second one aims at maximizing the sum rate. Two values are given for the searches of the maximum or minimum. The first value represents the number of repetitions, the second value denotes the size of the set in which the extremum is searched.

Table 6.4. Computational complexity per frame of the all-adaptive instance for the BS, where $\mathcal{R}_0 = 10$.

Adaptive algorithm	Multiplications/ divisions	Additions/ subtractions	Search maximum/ minimum (repetitions, size of set)
Design of grids of beams	348	111194	$(576, 12)$ $(576, 64)$
1. Allocation of resource blocks	191	192	$(192, 192)$; $(180, 12)$
1. Allocation of power and bits	84671	39744	$(1728, 12)$; $(1728, 64)$
2. Allocation of resource blocks	768	384	$(192, 12)$;
2. Allocation of power and bits	86400	38016	$(1728, 192)$;
Allocation of slots bits	396	297	$(99, 3)$; $(198, 2)$

Table 6.5. Computational complexity per frame of the all-adaptive instance for RS t, where $\mathcal{R}_t = 5$.

Adaptive algorithm	Multiplications/ divisions	Additions/ subtractions	Search maximum/ minimum (repetitions, size of set)
1. Allocation of resource blocks	191	192	$(192, 192)$; $(187, 5)$
1. Allocation of power and bits	84671	39744	$(1728, 5)$; $(1728, 64)$
2. Allocation of resource blocks	0	0	$(192, 5)$;
2. Allocation of power and bits	82944	38016	$(1728, 192)$;

The computational complexity given in both tables must be provided once per frame. In both tables, the comparison between the algorithms show that adaptive algorithms for the allocation of power and bits and for the design of grids of beams requires a large computational complexity. The computational complexity of the allocation of resource blocks and of the allocation of slots is rather low.

Since each AP allocates resource blocks, power and bits in two subframes, the computational complexity is roughly doubled compared to computational complexity given in Table 6.2 and in Table 6.3 and required for the all-adaptive instance if the objective of maximizing the minimum user rate is addressed. However, the example given in Section 6.2.2.1 shows that even if the computational complexity is doubled, the computational complexity is feasible for today's processors.

6.2.3 Design of Grids of Beams

6.2.3.1 Maximizing the Minimum User Rate

In this section, the impact of the adaptive algorithm for the design of grids of beams is evaluated. The distributed concept for orthogonal medium access is applied while the objective is the maximization of the minimum user rate in this section.

In order to show its impact, the adaptive algorithm is compared to the non-adaptive one defined in Section 4.2.2. Besides the all-adaptive and non-adaptive instance, the instance is considered in which the allocation of resource blocks, allocation of power and bits and the allocation of slots is solved adaptively, but the design of grids of beams is solved by the non-adaptive algorithm.

In Fig. 6.10, the average minimum user rate depending on the number of beams applied in the grids of beams is illustrated. The number of beams is the same for all APs in a snapshot. The all-adaptive instance and the non-adaptive instance are denoted as "all-adapt." and "non-adapt.", respectively. The instance using the non-adaptive algorithm for the design of grids of beams is denoted by indicating the subproblems solved by the BS and RSs adaptively and non-adaptively. For the design of Grids of Beams, the allocation of Resource Blocks, the allocation of power and bits and the allocation of slots the abbreviations "RB", "power" and "slots" are used. If a line is drawn over an abbreviation, the corresponding subproblem is solved by a non-adaptive algorithm, otherwise it is solved by an adaptive one. The results reveal that both instances lead to a huge gain compared to the non-adaptive

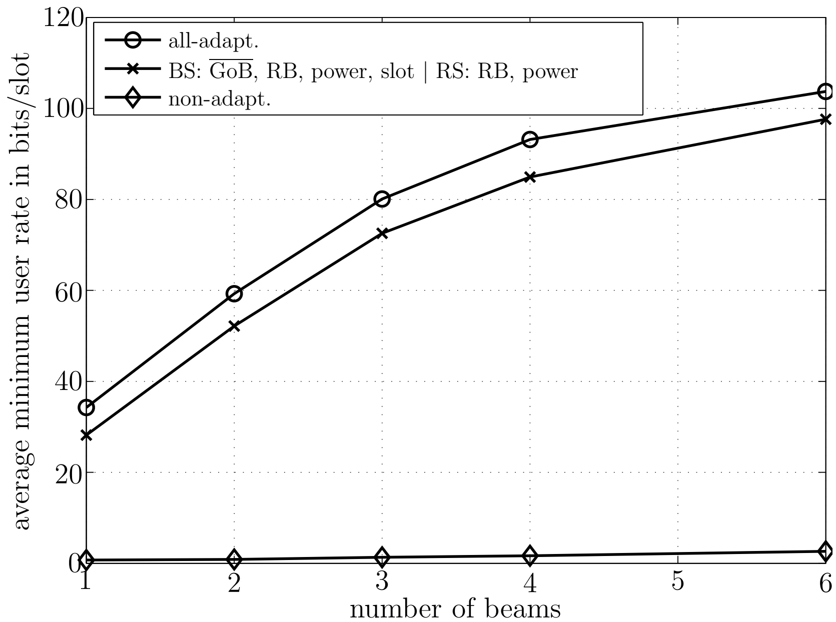

Figure 6.10. Comparison of the adaptive and non-adaptive algorithms for the design of grids of beams, $N_{\mathrm{UE}} = 20$.

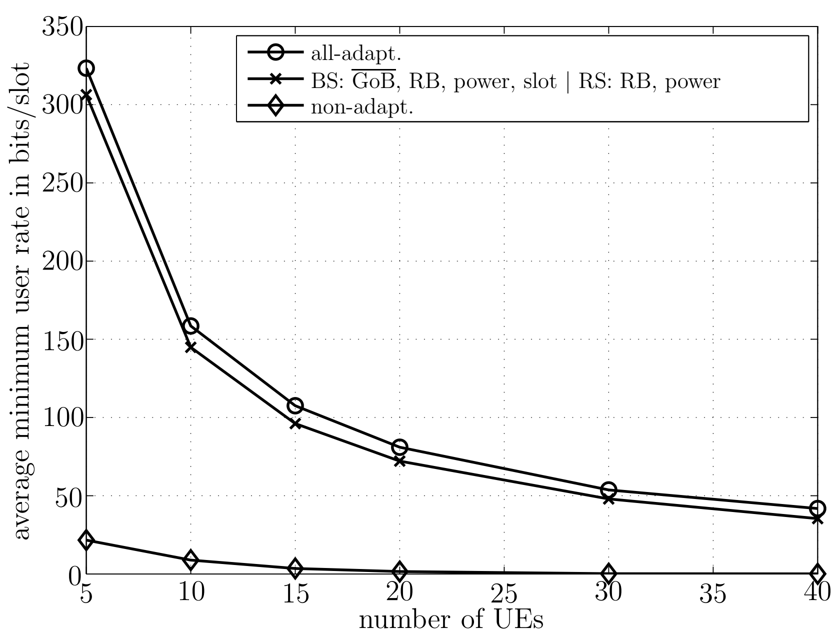

Figure 6.11. Comparison of the adaptive and non-adaptive algorithms for the design of grids of beams, $G_t = 3$ for all APs.

instance. The adaptive algorithm for the design of grids of beams always outperforms the non-adaptive one since beams are directed to the RSs and the UEs more often due to the utilization of the PDF representing the distribution of the UEs in the cell.

In Fig. 6.11, the dependency on the number of UEs is shown. The number of beams is three. A gain compared to the non-adaptive instance and compared to the instance using the non-adaptive algorithm for the design of grids of beams is achieved. The gain in absolute values of the average minimum user rate is rather constant. Since the average minimum user rate is a decreasing function of the number of UEs, the relative gain is larger if more UEs are in the cell. The reason is found in the definition of the demand values in (4.2) and (4.3). The demand values are determined according to the PDF representing the distribution of the UEs in the cell. The instantaneous distribution of the UEs converges to the PDF if the number of UEs is large.

In total, two conclusions are drawn. Firstly, the adaptive algorithm for the design of grids of beams leads to an improvement in terms of average minimum user rate. Secondly, it is not essential in order to outperform the non-adaptive instance if the adaptive or non-adaptive algorithm for the design of grids of beams is applied.

6.2.3.2 Maximizing the Sum Rate

In this section, the impact of the adaptive algorithm for the design of grids of beams is evaluated if the distributed concept for orthogonal medium access is applied while the objective is the maximization of the sum rate in this section.

In Fig. 6.12, the average sum rate is illustrated as a function of the number of beams applied in the grids of beams. Results of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.) are shown. The results of the instance using the non-adaptive algorithm for the design of grids of beams is also given. The notation of this instances is according to the one introduced in Section 6.2.3.1: The subproblems solved by the BS and RSs are denoted by abbreviations, a line drawn over an abbreviation indicates a non-adaptive algorithm and a missing line indicates an adaptive algorithm. The number of beams is the same for each AP. The minimum user rate is nearly always provided except for the non-adaptive instance which cannot provide the minimum user rate. The results reveal that the adaptive algorithm for the design of grids of beams does not improve the average sum rate compared to the non-adaptive one if the relative frequency of outages is close to zero. The reason is that the average sum rate is affected mainly by a low number of UEs close to the BS. These so called best UEs benefit from a low path

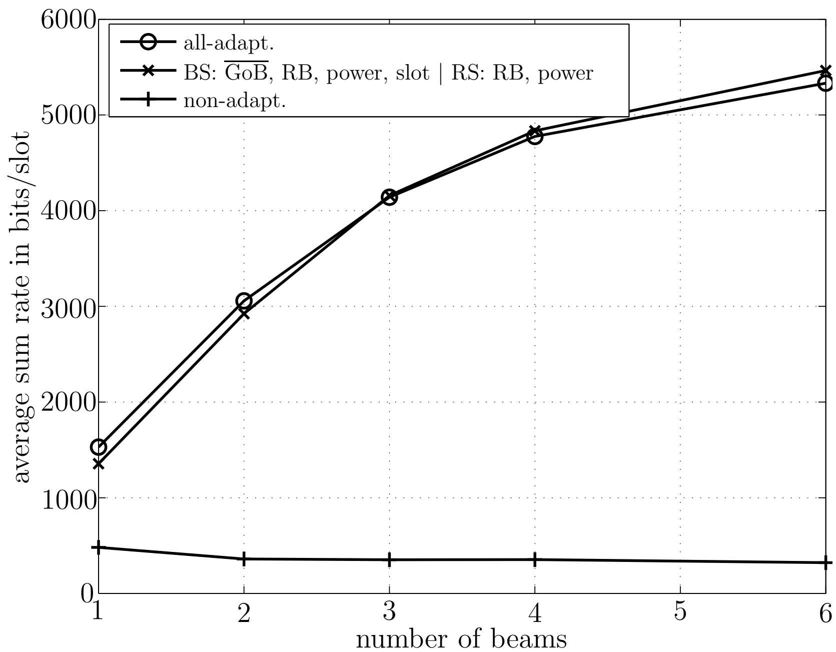

Figure 6.12. Comparison of the adaptive and non-adaptive algorithms for the design of grids of beams, $N_{\mathrm{UE}} = 20$, $R_{\min} = 10$ bits/slot.

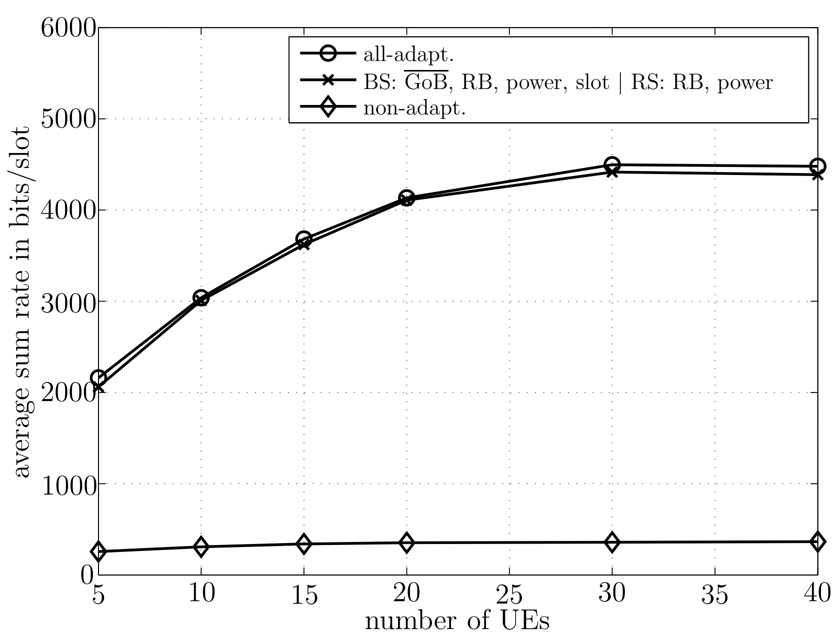

Figure 6.13. Comparison of the adaptive and non-adaptive algorithms for the design of grids of beams, $G_t = 3$ for all APs, $R_{\min} = 10$ bits/slot.

loss. In the adaptive algorithm for the design of grids of beams, this effect is not considered. The demand values are defined in (4.2) and (4.3) based on the PDF representing the distribution of the UEs. As already observed in Fig. 6.11, the adaptive algorithm performs well if the objective is affected by a large number of UEs. However, this is not the case if the objective is the maximization of the sum rate. The results would be different if more accurate information were available about the positions of the best UEs and the demand values would be adapted to this information.

In Fig. 6.13, the dependency on the number of UEs is shown. The number of UEs and the minimum user rates are chosen such that the relative frequency of outages tends to zero except for the non-adaptive instance. The previous result is approved that the adaptive algorithm has no impact on the average sum rate for a low relative frequency of outages.

In Fig. 6.14, the average sum rate is illustrated as a function of the minimum user rate. The results are only shown as long as the relative frequency of outages is less than 5%. If the relative frequency of outages is larger than 5%, the illustration of the average sum rate becomes too optimistic since only snapshots are considered for which the minimum user rate is achieved while the other snapshots are neglected. The adaptive algorithm is able to provide a higher minimum user rate per UE than the non-adaptive algorithm for the design of grids of beams. This becomes clear if the relative frequency of outages is evaluated.

In Fig. 6.15, the relative frequency of outages depending on the minimum user rate is shown. The result expected from Fig. 6.14 is approved. The result is explained as follows: Two subframes are allocated to each AP. The first one serves for achieving the minimum user rate and gets a stronger impact for large minimum user rates while the impact of the second one tends to zero. Since the allocation of resource blocks and allocation of power and bits is solved for this subframe with the objective of maximizing the minimum user rate, all UEs and not only the best one affects the average sum rate. As in Section 6.2.3.1, the adaptive algorithm for the design of grids of beams outperforms the non-adaptive one under these circumstances. The non-adaptive instance is not depicted in Fig. 6.15 since it always leads to an relative frequency of outages close to one.

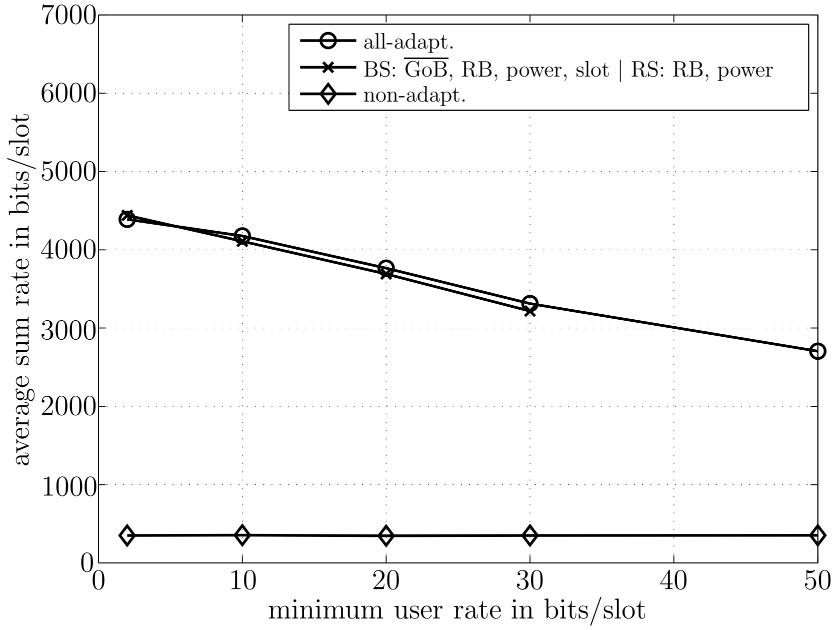

Figure 6.14. Comparison of the adaptive and non-adaptive algorithms for the design of grids of beams, $G_t = 3$ for all APs, $N_{\mathrm{UE}} = 20$.

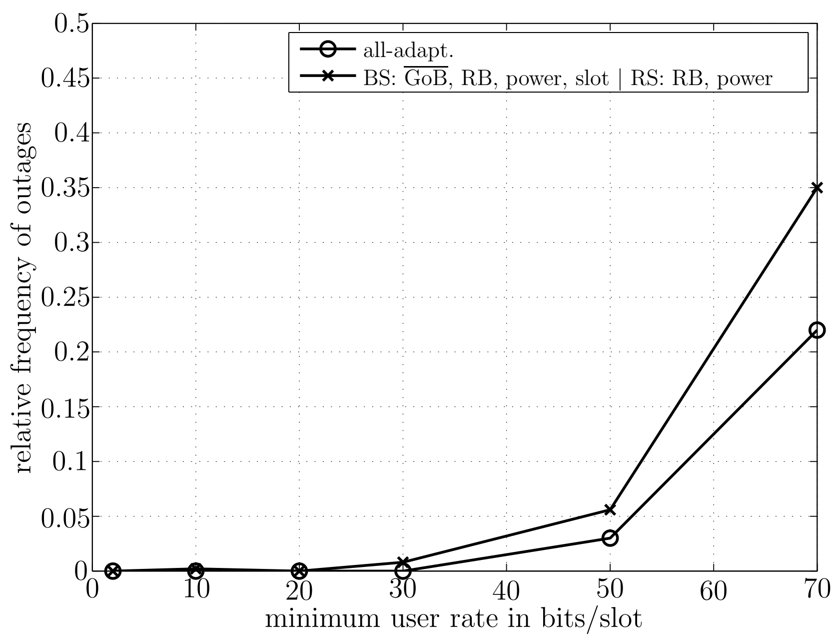

Figure 6.15. Relative frequency of outages of the adaptive and non-adaptive algorithms for the design of grids of beams, $G_t = 3$ for all APs, $N_{\mathrm{UE}} = 20$.

6.2.4 Allocation of Resource Blocks, Power and Bits

6.2.4.1 Maximizing the Minimum User Rate

In this section, the allocation of resource blocks and allcoation of power and bits is evaluated if the distributed concept for orthogonal medium access is applied with the objective of the maximization of the minimum user rate. In order to show the impact of the adaptive algorithms for the allocation of resource blocks and the allocation of power and bits, the adaptive algorithms are compared to the corresponding non-adaptive ones.

In Fig. 6.16, the average minimum user rate is depicted for various numbers of beams applied in the grids of beams. The number of beams applied is the same for each AP per snapshot. The abbreviations in the legend show which subproblems are solved by the BS or RS: design of Grids of Beams (GoB), allocation of Resource Blocks (RB), allocation of power and bits (power) and allocation of slots (slots). If a line is drawn over an abbreviation, the corresponding subproblem is solved by a non-adaptive algorithm, otherwise it is solved by an adaptive one. Furthermore, the results of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.) are shown. Actually, 2^4 instances defined by the combinations of algorithms can be evaluated if the BS may apply adaptive or non-adaptive algorithms and if the RSs may apply adaptive or non-adaptive algorithms for the allocation of resource blocks and the allocation of power and bits. Here, a selection of instances is evaluated. The selection is motivated by replacing successively one adaptive algorithm by a non-adaptive one. Since the relays shall be rather cheap compared to the BS according to Section 1.1, the computational complexity is reduced at the RSs at first.

In the following, the considered instances are compared successively. The best performing one is the all-adaptive instance. The comparison to the one in which the RS solves the allocation of power and bits non-adaptively and the other subproblems adaptively shows the impact of the adaptive algorithm for allocating power and bits by the RS. Especially, if the number of applied beams is large, the adaptive algorithm leads to an improvement. Since more beams are applied, stronger inter-beam interference occurs and is successfully mitigated by the adaptive algorithm for allocating power and bits, but is not mitigated by the non-adaptive one. If the allocation of resource blocks and the allocation of power and bits are solved by the RS with the non-adaptive algorithms, the average minimum user rate drops dramatically. Links use resource blocks, which are not suitable for them. Resource blocks are allocated to RS-to-UE links although beams are not directed to the receiving UEs. For large number of beams, the average minimum user rate even decreases since resource blocks

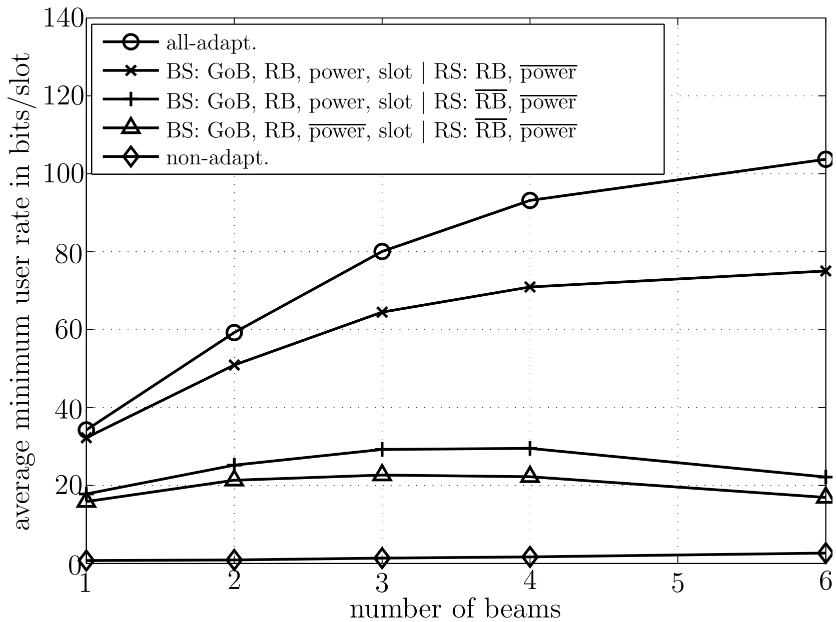

Figure 6.16. Comparison of various instances, $N_{\text{UE}} = 20$.

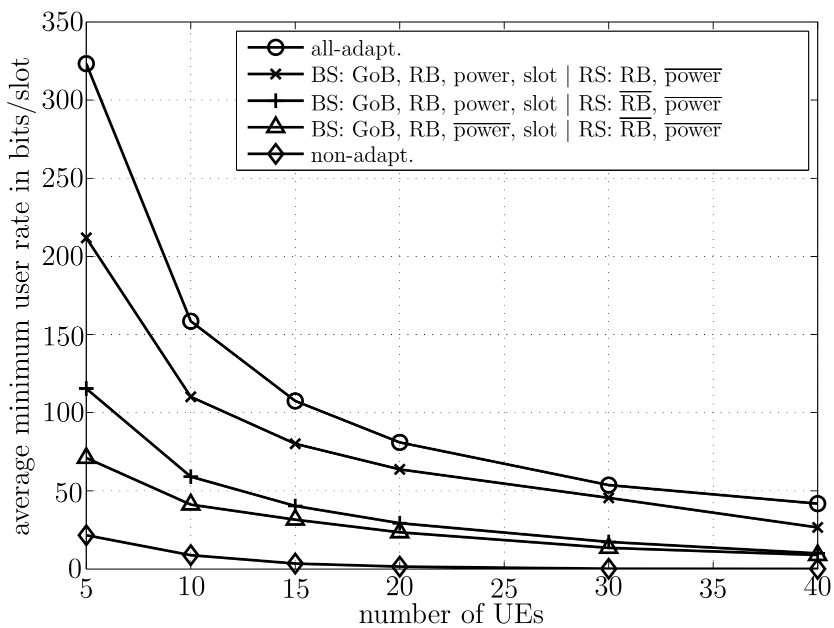

Figure 6.17. Comparison of various instances, $G_t = 3$ for all APs.

are allocated to RS-to-UE links although the receivers suffer from strong inter-beam interference. In total, it is concluded that the adaptive allocation of resource blocks by the RSs has a stronger impact than the adaptive allocation of power and bits. The same conclusion is drawn from the results if the adaptive algorithms concerning the allocation of resource blocks, power and bits and applied by the BS are replaced by the non-adaptive ones. Besides the stronger impact, it must be stated that the adaptive algorithms for the allocation of resource blocks require much less computational complexity than the adaptive algorithms for the allocation of power and bits as shown in Table 6.2 and Table 6.3. In total, it is concluded that an adaptive allocation of resource blocks is essential for all APs due to the performance gains and the low computational complexity. The adaptive allocation of power and bits leads to an improvement of the average minimum user rate, but requires a large computational complexity.

In Fig. 6.17, the average minimum user rate is given as a function of the number of UEs. The conclusions drawn from the results of Fig. 6.16 are approved.

6.2.4.2 Maximizing the Sum Rate

In this section, the allocation of resource blocks and allcoation of power and bits is evaluated if the distributed concept for orthogonal medium access is applied with the objective of the maximization of the sum rate. The instances considered in this sections result from the same combinations of adaptive and non-adaptive algorithms as in Section 6.2.4.1, but the algorithms are changed with respect to the objective of the maximization of the sum rate.

In Fig. 6.18, the average sum rate is given as a function of the number of beams applied in the grids of beams. Only the snapshots in which each user achieves the minimum user rate are considered and the results are only plotted if the relative frequency of outages is below 5%. These two rules are not applied to the non-adaptive instance which acts only as a reference. As expected, the best performing instance is the all-adaptive one. If the RSs apply the non-adaptive algorithm for the allocation of power and bits and the other subproblems are solved adaptively, the average sum rate keeps nearly unchanged. For both instances, a RS is allocated typically only one slot in order to fulfill the minimum data rate for all UEs using a two-hop connection, but the average sum rate is affected mainly by the UEs having established direct connections. If the RSs apply also the non-adaptive algorithm for the allocation of resource blocks, the average sum rate is reduced. Resource blocks using beams not directed to the UEs are allocated to the RS-to-UE links. More slots are required to

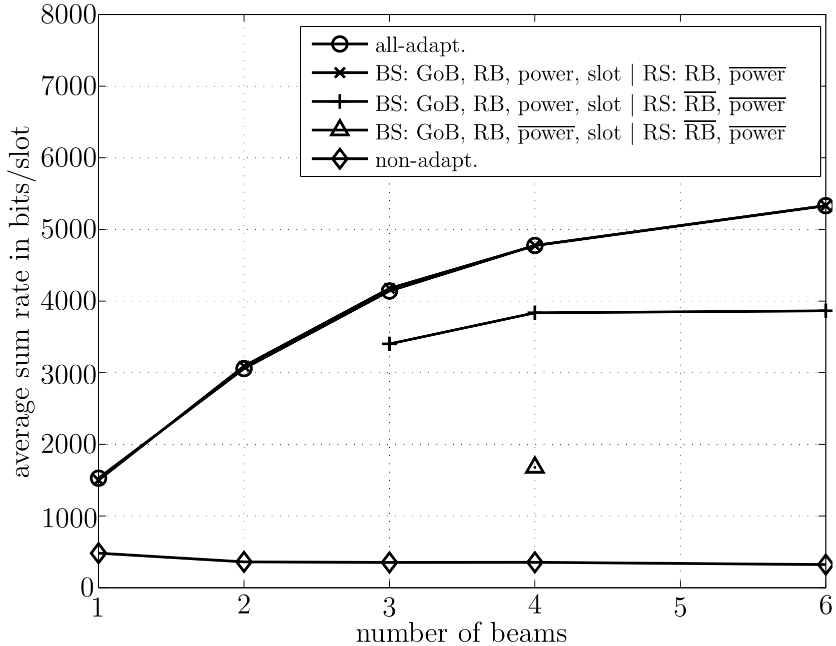

Figure 6.18. Comparison of various instance, $N_{\mathrm{UE}} = 20$, $R_{\min} = 10$ bits/slot.

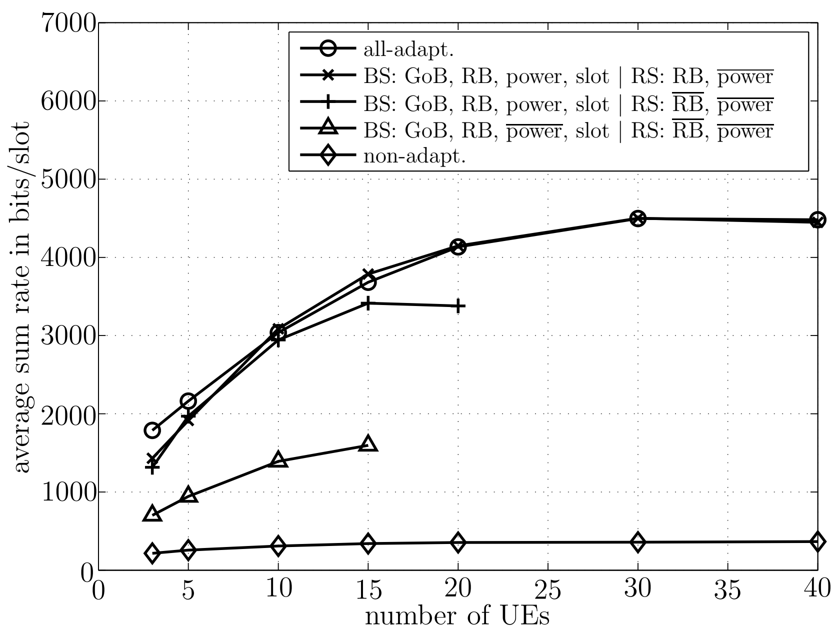

Figure 6.19. Comparison of various instances, $G_t = 3$ for all APs, $R_{\min} = 10$ bits/slot.

offer the minimum user rate to the UEs having established two-hop connections. If only one beam is applied per time-frequency unit, an relative frequency of outages larger than 5% cannot be provided. The average sum rate drops dramatically if the BS applies the non-adaptive algorithm for the allocation of power and bits. This strong impact approves further that the average sum rate is affected mainly by the UEs having established direct connections. Additionally, the user rate of the UEs having established direct connections decreases such that the relative frequency of outages is larger than 5% except for three beams.

In Fig. 6.19, the average sum rate depending on the number of UEs is depicted. Again, results are only shown if the relative frequency of outages is below 5% except for the non-adaptive instance. The comparison of the all-adaptive instance and the instance in which the allocation of power and bits solved by the RSs is the only non-adaptive algorithm shows that the adaptive algorithm applied by the RSs for the allocation of power and bits improves the average sum rate for small number of UEs. If only a small number of UEs exists, the UEs having established two-hop connections have a stronger impact on the average sum rate since a probability significantly larger than zero exists that all UEs served by a direct connection suffer from a weak channel gain. The results reveal that the minimum user rate can only be provided with non-adaptive algorithms for the allocation of resource blocks by the RS and the allocation of power and bits by the BS if the number of UEs is rather small. This effect occurs since the sum of minimum user rate values increases with more UEs. This effects becomes clearer in the next figure.

In Fig. 6.20, the relative frequency of outages depending on the minimum user rate is illustrated. The best performance in terms of relative frequency of outages is achieved by the all-adaptive instance. If the RSs apply the non-adaptive algorithm for the allocation of power and bits, a first degradation of the relative frequency of outages is revealed. The relative frequency of outages increases dramatically if the RSs also apply the non-adaptive allocation of resource blocks. Whether the BS applies the adaptive or non-adaptive algorithm for the allocation of power and bits, has only a small impact on the relative frequency of outages. Since the objective of the allocation of resource blocks, power and bits in the first subframe allocated to an AP is the maximization of the minimum user rate, these results are in line with the conclusions drawn in Section 6.2.4.1: An adaptive allocation of resource blocks is essential for all APs, while the adaptive allocation of power and bits leads to an improvement.

In order to summarize the results of this section, two conclusions are drawn. At first, the allocation of resource blocks must be solved by all APs with the adaptive algorithm in order to achieve a low relative frequency of outages. At second, the adaptive algorithm for the allocation of power and bits by the BS has a strong impact on the average sum rate, but the one applied by the RS has not. However, the adaptive algorithms for the allocation of power and bits have a larger computational complexity than the adaptive algorithms for the allocation of resource blocks according to Table 6.4 and Table 6.5.

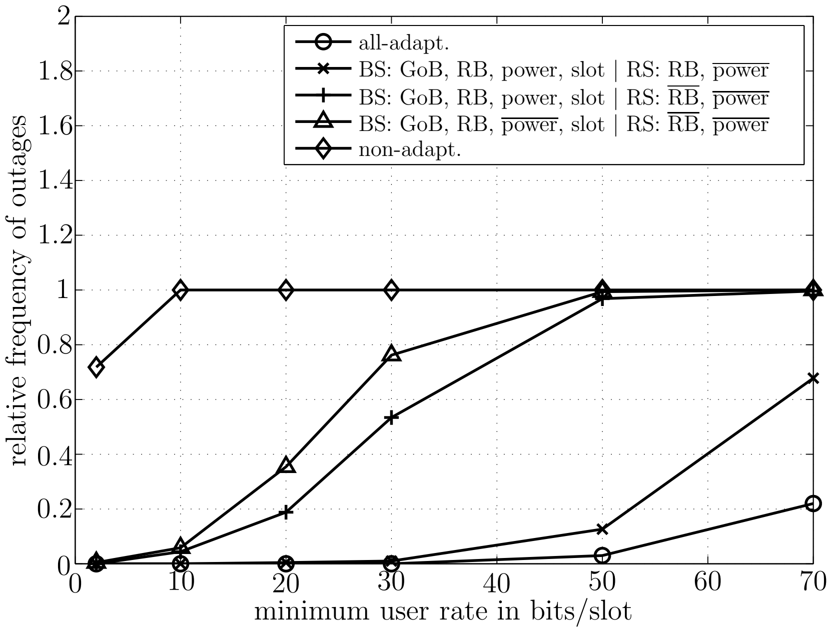

Figure 6.20. Relative frequency of outages of various instances, $N_{\mathrm{UE}} = 20$, $G_t = 3$ for all APs.

6.2.5 Allocation of Slots

6.2.5.1 Maximizing the Minimum User Rate

In this section, the allocation of slots is evaluated if the distributed concept for orthogonal medium access is applied with the objective of the maximization of the minimum user rate. In order to show the impact of the adaptive algorithm for the allocation of slots, the adaptive algorithm is compared to the corresponding non-adaptive one introduced in Section 5.2. Since the non-adaptive algorithm is only defined in the distributed concept for reuse medium access, this concept is applied for the comparison although an orthogonal medium access is considered.

The comparison between the adaptive and non-adaptive algorithm is mainly affected by the number of UEs in the cell. In Fig. 6.21, the average minimum user rate depending on the number of UEs is given. The all-adaptive and non-adaptive instances are given as references. Additionally two instances are given serving for the actual comparison between the adaptive and non-adaptive algorithm for the allocation of slots. The abbreviations used in the legend indicate the subproblems solved by the BS or RS: design of Grids of Beams (GoB), allocation of Resource Blocks (RB) and allocation of slots (slot). A line drawn over an abbreviation denotes a non-adaptive algorithm. If the adaptive algorithm is applied for the allocation of slots, the non-adaptive algorithm is applied for the allocation of power and bits. Hence, gains are only achieved by the adaptive algorithm for the allocation of slots. The comparison reveals that the adaptive allocation of slots increases the minimum user rate especially for small number of UEs. The gain decreases if more UEs are in the cell. If the number of UEs is quite low, the number of UEs served by an AP fluctuates strongly. Hence, the number of slots allocated to the APs also fluctuate strongly in order to balance the user rates within the cell. However, this effect vanishes for a very large number of UEs. The instances using the adaptive and non-adaptive algorithm for the allocation of slots outperform the non-adaptive instance clearly. Hence, it is not essential in order to outperform the non-adaptive instance if the adaptive or non-adaptive algorithm for the allocation of slots is applied.

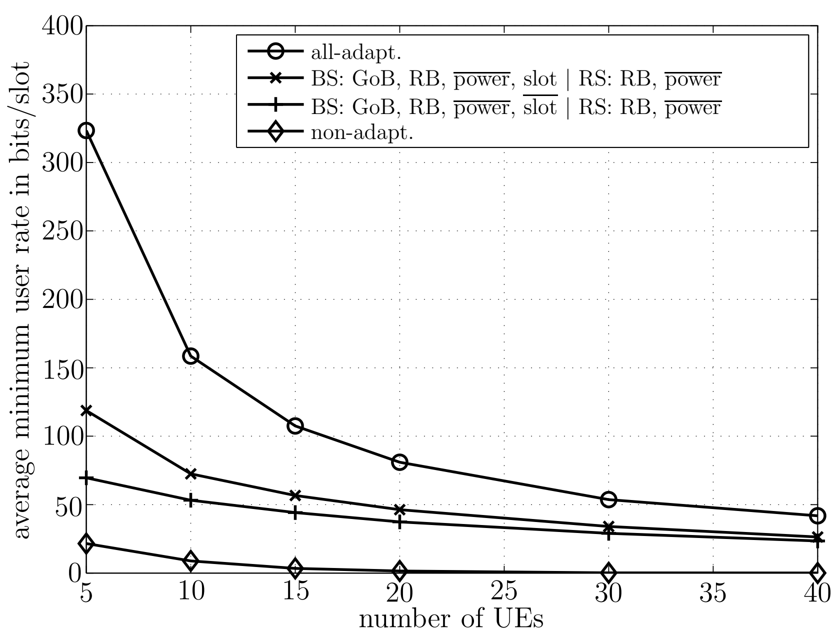

Figure 6.21. Comparison of various instances to show the influence of the adaptive algorithm for the allocation of slots, $G_t = 3$ for all APs.

6.2.5.2 Maximizing the Sum Rate

In this section, the allocation of slots is evaluated if the distributed concept for orthogonal medium access is applied with the objective of the maximization of the sum rate. A comparison is made between the adaptive algorithm and the non-adaptive one. Since the non-adaptive one is only defined in the distributed concept for reuse medium access, this concept is applied for the comparison.

In Fig. 6.22, the average sum rate as a function of the number of UEs is illustrated. The abbreviations used in the legend show again which subproblems are solved non-adaptively and adaptively by the BS or RS. The all-adaptive and non-adaptive instances are given as references. Only the snapshots are considered in which the minimum user rate is achieved by all UEs, except for the non-adaptive concept. The results are only shown if the relative frequency of outages is less than 5%. The performance of the adaptive algorithm for the allocation of slots is revealed if

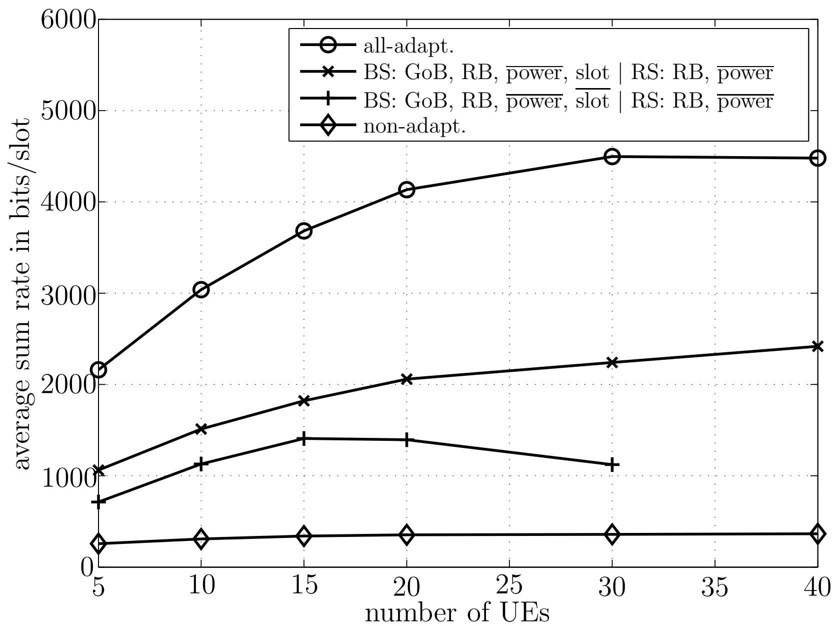

Figure 6.22. Comparison of various instances to show the influence of the adaptive algorithm for the allocation of slots, $G_t = 3$ for all APs, $R_{\min} = 10$ bits/slot.

the instance using the non-adaptive algorithms for the allocation of power and bits and the adaptive algorithm for the allocation of slots is compared to the instance using the non-adaptive algorithms for the allocation of power and bits and for the allocation of slots. The comparison shows that the adaptive algorithm leads to two advantages. At first, the adaptive algorithm leads to an relative frequency of outages smaller than 5% even for 40 UEs where each one has a minimum rate of 10 bits/slot. The non-adaptive one cannot provide this. If more than 30 UEs are in the cell, the average sum rate is not depicted any longer for the corresponding instance since the relative frequency of outages exceeds 5%. At second, the average sum rate is improved. For 5 UEs until 15 UEs, the absolute gain in terms of average sum rate is nearly constant. Hence, the relative gain is large especially for small number of UEs as already shown in Section 6.2.5.1, where the objective is the maximization of the minimum user rate. For larger number of UEs, the average sum rate achieved by the instance using the non-adaptive algorithms for the allocation of power and bits and for the allocation of slots decreases. The APs must allocate more resource blocks to links in order to achieve the minimum user rate and not in order to maximize the sum rate. This is avoided if the adaptive algorithm for the allocation of slots is applied.

6.2.6 Summary for the Distributed Concept for Orthogonal Medium Access

In Section 6.2, the following results are derived:

- The distributed concept for orthogonal medium access enables adaptive algorithms concerning the design of grids of beams, the allocation of resource blocks, power and bits and the allocation of slots. The subproblems are solved based on a signalling overhead from RS to BS which is much lower than required for a solution at a central unit of the cell.

- If the objective is defined by (P1), the following gains compared to non-adaptive algorithms are achieved:

  - The adaptive algorithm for the design of grids of beams improves the average minimum user rate slightly.
  - The adaptive algorithms applied by the BS and RSs for the allocation of resource blocks achieve huge gains in terms of average minimum user rate. Their impact is the strongest among the adaptive algorithms and only a low computational complexity is required compared to the other adaptive algorithms.
  - The adaptive algorithms applied by the BS and RSs for the allocation of power and bits improves the average minimum user rate further. The gains are smaller than the ones achieved by the adaptive algorithms for the allocation of resource blocks. The computational complexity is larger than the one for the adaptive algorithms for the allocation of resource blocks.

  - The adaptive algorithm for the allocation of slots leads to a significant improvement for a small number of UEs in the cell.

- If the objective is defined by (P2), the following gains compared to non-adaptive algorithms are achieved:

  - The adaptive algorithm for the the design of grids of beams improves the relative frequency of outages while the average sum rate is nearly the same.
  - The adaptive algorithms applied by the BS and RSs for the allocation of resource blocks achieve huge gains in terms of relative frequency of outages. Only a low computational complexity is required related to the other adaptive algorithms.
  - The adaptive algorithm for the allocation of power and bits improve the average sum rate and the relative frequency of outages if applied by the BS. The adaptive algorithm applied by the RS is important to guarantee the minimum user rate, but not for optimizing the average sum rate. The required computational complexity is larger than the one for the adaptive algorithms for the allocation of resource blocks.
  - The adaptive algorithm for the the allocation of slots leads to a gain in terms of relative frequency of outages and average sum rate.

6.3 Distributed Concept for Reuse Medium Access

6.3.1 Evaluation Scenario

The distributed concept for reuse medium access is also evaluated by snapshot simulations. The scenario is the same one as defined in Section 6.2.1. The only exception is related to the medium access of the APs. In this Section 6.3, it is assumed that a frame consists of two subframes. The BS transmits in the first subframe and both RSs transmit in the second one simultaneously. Since the allocation of slots is always solved by the non-adaptive algorithm, the sizes of the two subframes are fixed for a particular combinations of the number of beams applied by the BS and RSs in the grids of beams. The sizes of the first subframe determined according to (5.4) as a function of the beams applied by the BS and the RSs for a time-frequency unit are listed in Tab. 6.6.

Table 6.6. Subframe sizes for different numbers of applied beams in a grid of beams.

	Beams at RS				
Beams at BS	1	2	3	4	6
1	69	78	83	86	89
2	63	69	75	78	83
3	59	65	69	73	78
4	57	62	66	69	74
6	55	59	62	65	69

6.3.2 Concept Evaluation

6.3.2.1 Maximizing the Minimum User Rate

In this section, the distributed concept for reuse medium access is evaluated if the objective of the maximization of the minimum user rate is addressed according to problem (P1). At first, the concept is evaluated concerning the signalling overhead, then concerning the average minimum user rate. Finally, the computational complexity is given for the considered concept.

The signalling overhead required for an optimum, central solution is given by (6.2). Since the channel gain values of each link between a UE served by an RS and the interfering RSs are required, this signalling overhead is larger than the signalling overhead required for an central solution for orthogonal medium access given by (6.4). For the distributed concept for reuse medium access, the signalling overhead keeps the same as defined in (6.5) for the distributed concept for orthogonal medium access. Hence, the conclusions drawn from the results depicted in Fig. 6.2 and Fig. 6.3 are also valid here, i.e., the proposed concept requires less signalling overhead than the optimum, central solution and the central solution is impracticable.

In the following, the performance of the proposed concept is evaluated in terms of average minimum user rate. The all-adaptive and non adaptive instance are considered. The all-adaptive instance is defined by the combination of the three adaptive algorithms for the design of grids of beams, the allocation of resource blocks by the RSs and the allocation of resource blocks by the BS. The non-adaptive instance is given by the combination of the non-adaptive algorithms. In Fig. 6.23, the average minimum user rate is depicted depending on the number of beams applied by the BS. The all-adaptive instance (all-adapt.) is compared to the non-adaptive one (non-adapt.) in which the RS applies always six beams. The results of the all-adaptive instance are given if the RSs apply one, three or six beams. The results

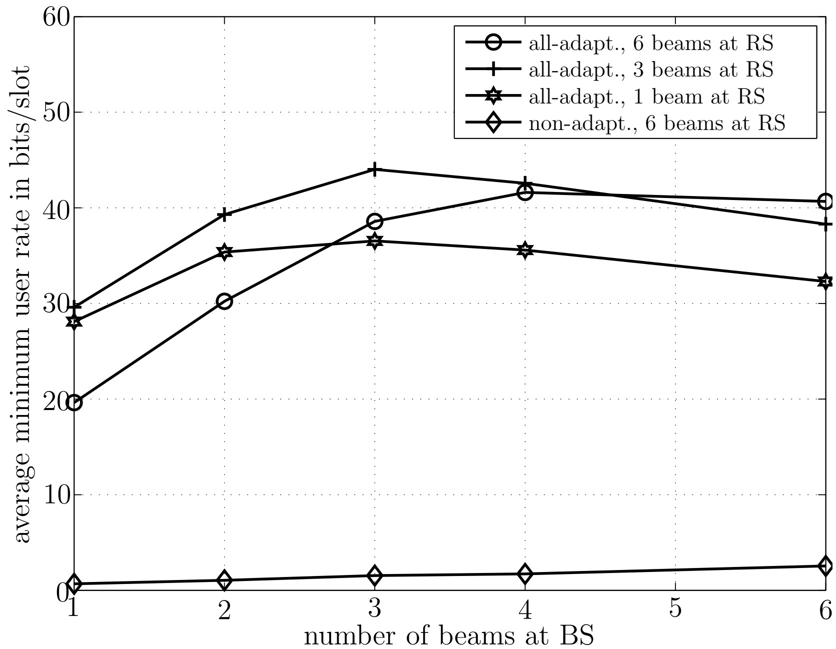

Figure 6.23. Average minimum user rate of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.) as a function of the number of beams used by the BS, various numbers of beams at the RSs for the all-adaptive instance, 6 beams at the RSs for the non-adaptive instance, $N_{\text{UE}} = 20$.

of the case in which the RSs apply two or four beams, are not depicted in order to give a clear representation. The explanation of the results of the all-adaptive instance as a function of beams applied by the BS starts with the results related to one beam applied by the RS: If the BS applies more than one beam, the performance is increased since more resource blocks are available for the BS-to-RS and BS-to-UE links. Simultaneously, the inter-beam interference becomes stronger. If more than three beams are applied by the BS, the inter-beam interference is so strong that the performance decreases with additional beams. If the RSs apply three beams instead of one, the performance of the all-adaptive instance is improved since more resource blocks are available for the RS-to-UE links. If the RSs apply six beams instead of three, the performance is decreased since the co-channel interference between the RSs and inter-beam interference become too strong. In order to find the number of beams applied by the APs such that the average minimum user rate is maximized, all combinations of applied number of beams must be tested. For the exemplary scenario assumed for the results of Fig. 6.23, the best combination is that all APs apply three beams. Even if the BS or RSs apply two or four beams, the average minimum user rate is not increased. For each combination of beams, the all-adaptive instance outperforms the non-adaptive one.

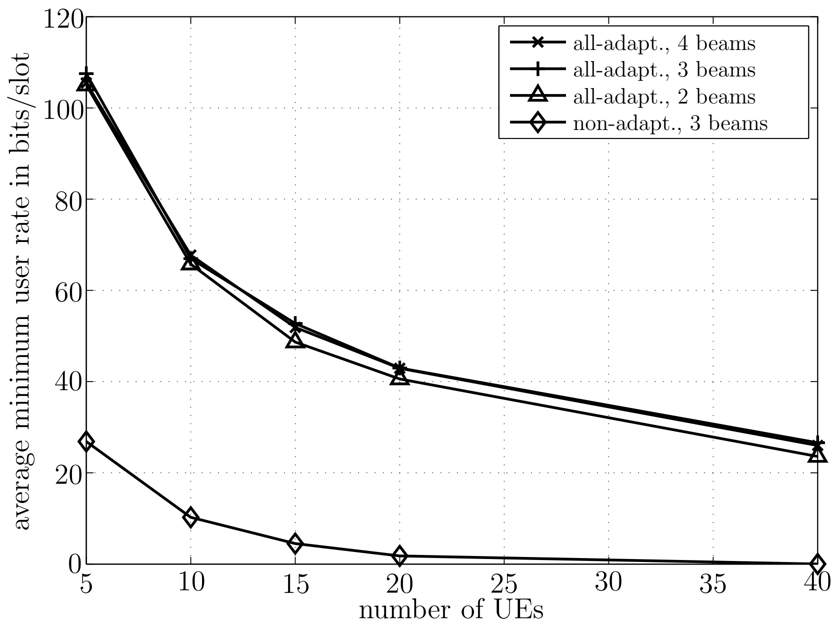

Figure 6.24. Average minimum user rate of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.), same number of beams at the BS and RSs, various numbers of beams for the all-adaptive instance, 3 beams for the non-adaptive instance.

In Fig. 6.24, the average minimum user rate is illustrated as a function of the number of UEs. The BS and the RSs apply two, three or four beams. The results show that more beams do not lead to a larger average minimum user rate since the gain caused by additional resource blocks is compensated by a stronger co-channel interference. However, the all-adaptive instance performs much better than the non-adaptive instance.

In contrast to the non-adaptive instance, the all-adaptive instance requires a significant computational complexity. The values for the computational complexity are given for the BS and an RS in Table 6.7 and Table 6.8, respectively if the all-adaptive instance is applied. The values are derived from Table 4.3 and Table 5.1. The number G_t of beams is given by the default value, i.e., $G_t = 3$ for all t. The assumption is made that ten UEs are assigned to the BS and five UEs to a RS. Both tables show that the adaptive algorithm for the design of grids of beams requires a large computational complexity compared to the adaptive algorithms for the allocation of resource blocks. However, the all-adaptive instance requires less computational complexity than it is the case if the orthogonal medium access is considered as approved by a comparison to Table 6.2 and to Table 6.3. As shown in the example given in Section 6.2.2.1, the computational complexity is feasible for today's processors.

Table 6.7. Computational complexity per frame of the all-adaptive instance for the BS, where $|\mathcal{R}_0| = 10$.

Adaptive algorithm	Multiplications/ divisions	Additions/ subtractions	Search maximum/ minimum (repetitions, size of set)
Design of grids of beams	348	86618	$(576, 12)$; $(576, 64)$
Allocation of resource blocks	191	192	$(192, 192)$; $(180, 12)$

Table 6.8. Computational complexity of the all-adaptive instance for RS t, where $|\mathcal{R}_t| = 5$.

Adaptive algorithm	Multiplications/ divisions	Additions/ subtractions	Search maximum/ minimum (repetitions, size of set)
Allocation of resource blocks	191	192	$(192, 192)$; $(187, 5)$

6.3.2.2 Maximizing the Sum Rate

In this section, the distributed concept for reuse medium access is evaluated if the objective is the maximization of the sum rate according to problem (P2). The concept is evaluated concerning the average sum rate and relative frequency of outages. An analysis of the signalling overhead is needless as the signalling overhead of the distributed concept for reuse medium access and of an optimum, central solution is unchanged if the objective is related to (P2) instead of (P1).

In Fig. 6.25, the average sum rate depending on the number of applied beams is given for the all-adaptive instance and the non-adaptive one. The average sum rate is measured only over the snapshots in which the minimum user rate is fulfilled by all UEs. This rule is not applied to the non-adaptive instance since the minimum user rate cannot be provided. The results related to the cases in which the BS applies one and the RS one and six beams, respectively, are not depicted since the relative frequency of outages is larger than 5%. In the former case, the number of resource blocks is not large enough in order to achieve a relative frequency of outages less than

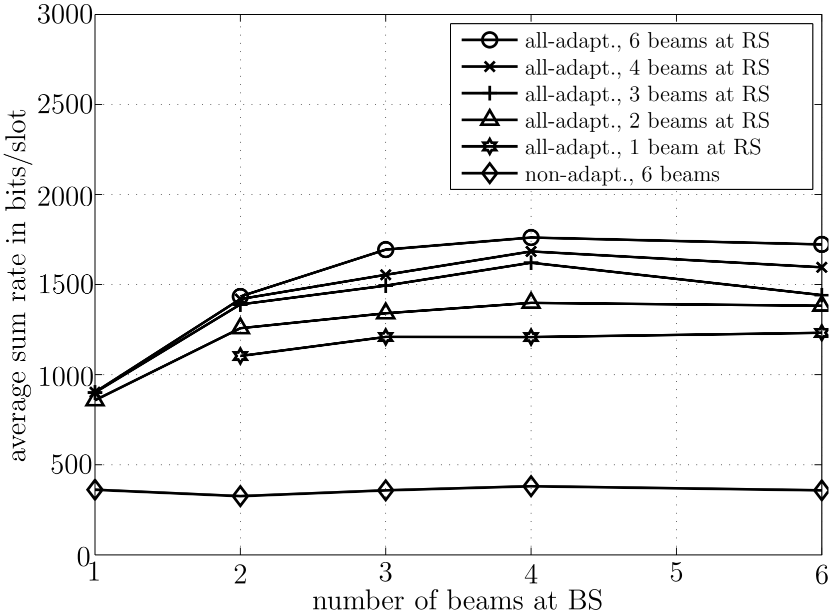

Figure 6.25. Average sum rate of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.) as a function of the number of beams used by the BS, various numbers of beams at the RSs for the all-adaptive instance, 6 beams at the RSs for the non-adaptive instance, $N_{\mathrm{UE}} = 20$, $R_{\min} = 10$ bits/slot.

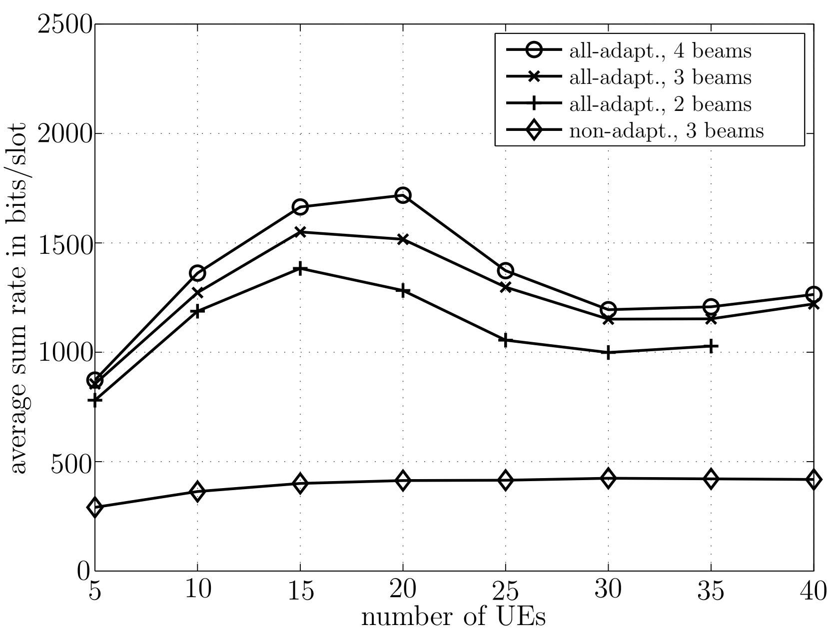

Figure 6.26. Average sum rate of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.), same number of beams at the BS and RSs, various numbers of beams for the all-adaptive instance, 3 beams for the non-adaptive instance, $R_{\min} = 10$ bits/slot.

5%. In the latter case, the interference is too strong. More applied beams lead to a higher average sum rate, but the gain is rather small for large number of beams. Results in which more than six beams are applied by an AP per time-frequency unit are not shown since a weak performance caused by strong inter-beam interference. In contrast to the results depicted in Fig. 6.23, where the objective of maximizing the minimum user rate is considered, the trend of the results in Fig. 6.25 is that beams applied by the APs leads to a further gain. Since the objective is the maximization of the sum rate, resource blocks are allocated to a link affected by low inter-beam interference. If the objective is the maximization of the minimum user rate, resource blocks need not to be allocated to a link achieving a low data rate.

In Fig. 6.26, the average sum rate is shown as a function of the number of UEs. The all-adaptive instance is considered if the APs apply two, three and four beams and the non-adaptive instance is considered if the APs apply three beams. The results of the all-adaptive instance reveal a large gain compared to the non-adaptive one. The characteristic of the function representing the result of a setup of the all-adaptive instance is well known from the results of Fig. 6.7: For small number of UEs, the multiuser diversity affect the results. For a larger number, the impact of the minimum user rate prevails. Finally, the average sum rate starts increasing since the relative frequency of outages becomes close to 5% and the average sum rate values become too optimistic since the snapshots counted as an outage are not considered in the average sum rate. Hence, the results of the all-adaptive instance are not presented if the relative frequency of outages becomes larger than 5%. It is revealed that each setup achieves this relative frequency of outages nearly for the same number of UEs.

In Fig. 6.27, the relative frequency of outages depending on the minimum user rate is illustrated. The all-adaptive and the non-adaptive instance are considered if the APs apply three beams. As already stated, the non-adaptive one cannot provide the minimum user rate in most of the cases even if the minimum user rate is rather small. Hence, the all-adaptive instance outperforms the non-adaptive one.

In contrast to the non-adaptive instance, the all-adaptive instance requires a significantly higher computational complexity. The values for the computational complexity are given for the BS and an RS in Table 6.9 and Table 6.10, respectively if the all-adaptive instance is applied. The values are derived form Table 5.1 and Table 5.2. The number G_t of beams is three for all t. Ten UEs are assigned to the BS and five UEs to a RS. Both tables show that the adaptive algorithm for the design of grids of beams requires a large computational complexity compared to the adaptive algorithms for the allocation of resource blocks. A comparison to Table 6.7 and Table 6.8 shows that the

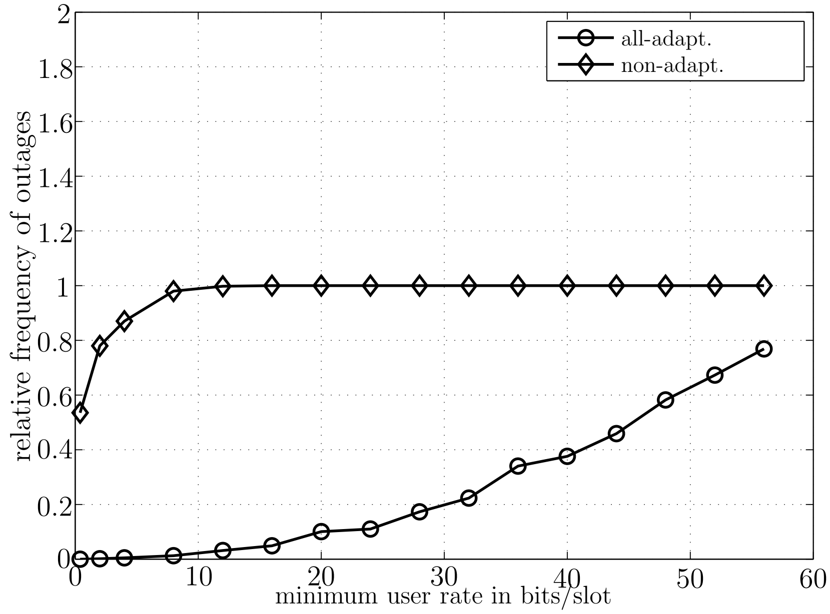

Figure 6.27. Relative frequency of outages of the all-adaptive instance (all-adapt.) and the non-adaptive instance (non-adapt.), $G_t = 3$ for all APs, $N_{\mathrm{UE}} = 20$.

computational complexity is nearly the same as the one required if the objective is the maximization of the minimum user rate.

Table 6.9. Computational complexity of the all-adaptive instance for the BS, where $|\mathcal{R}_0| = 10$.

Adaptive algorithm	Multiplications/ divisions	Additions/ subtractions	Search maximum/ minimum (repetitions, size of set)
Design of grids of beams	348	86618	$(576, 12)$; $(576, 64)$
Allocation of resource blocks	192	180	$(192, 192)$; $(180, 12)$

Table 6.10. Computational complexity of the all-adaptive instance for RS t, where $|\mathcal{R}_t| = 5$.

Adaptive algorithm	Multiplications/ divisions	Additions/ subtractions	Search maximum/ minimum (repetitions, size of set)
Allocation of resource blocks	192	187	$(192, 192)$; $(187, 5)$

6.3.3 Design of Grids of Beams

6.3.3.1 Maximizing the Minimum User Rate

In this section, the impact of the adaptive algorithm for the design of grids of beams is analyzed. The objective is the maximization of the minimum user rate. In Fig. 6.28, the all-adaptive instance (all-adapt.) is compared to the non-adaptive instance (non-adapt.) and to the instance using the non-adaptive algorithm for the design of grids of beams. The latter instance is denoted in the figure by the APs solving the subproblems and by the abbreviations of the subproblems: design of Grids of Beams (GoB) and allocation of Resource Blocks (RB). A line is drawn over GoB in order to indicate that this subproblem is solved by the non-adaptive algorithm. The instances are compared in terms of average minimum user rate as a function of the number of beams applied by the APs. The graph of the all-adaptive instance is in line with results of Fig. 6.23. The average minimum user rate of the instance using the non-adaptive algorithm is not a smooth function of the number of beams. The co-channel interference affecting the RS-to-UE links depends on the beams applied by the RSs. These beams are chosen according to Section 4.2.2 without taking into account co-channel interference. The adaptive algorithm for the design of grids of beams outperforms the non-adaptive one except for the case of six beams. In the case of six beams, so many beams are applied that co-channel and inter-beam interference cannot be mitigated by the adaptive algorithm for the design of grids of beams better than mitigated by the non-adaptive algorithm. As one can expect by the broad beamwidth depicted in Fig. 4.1, a grid of beams is designed in terms of minimum inter-beam interference optimally if six out of twelve beams are chosen such that neighboring beams are not grouped together. This is achieved by the adaptive and the non-adaptive algorithm for the design of grids of beams. The antenna pattern resulting from six beams in a grid of beams is nearly omnidirectional. Hence, co-channel interference cannot be mitigated by the adaptive algorithm.

In Fig. 6.29, the average minimum user rate is depicted as a function of the number of UEs. The same instances are considered as in the previous figure. The adaptive algorithm for the design of grids of beams outperforms the non-adaptive one especially if the number of UEs is large. This effect is already observed in Fig. 6.11. The demand values are determined according to the PDF representing the distribution of the UEs in the cell. The instantaneous distribution of the UEs converges to the PDF if the number of UEs is large.

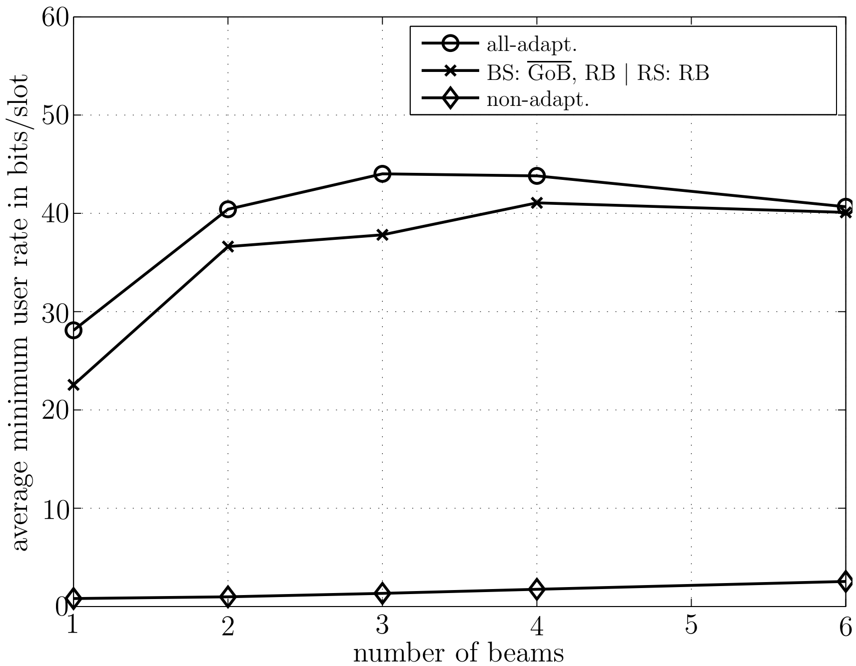

Figure 6.28. Comparison of the adaptive and non-adaptive algorithms for the design of grids of beams $N_{\mathrm{UE}} = 20$.

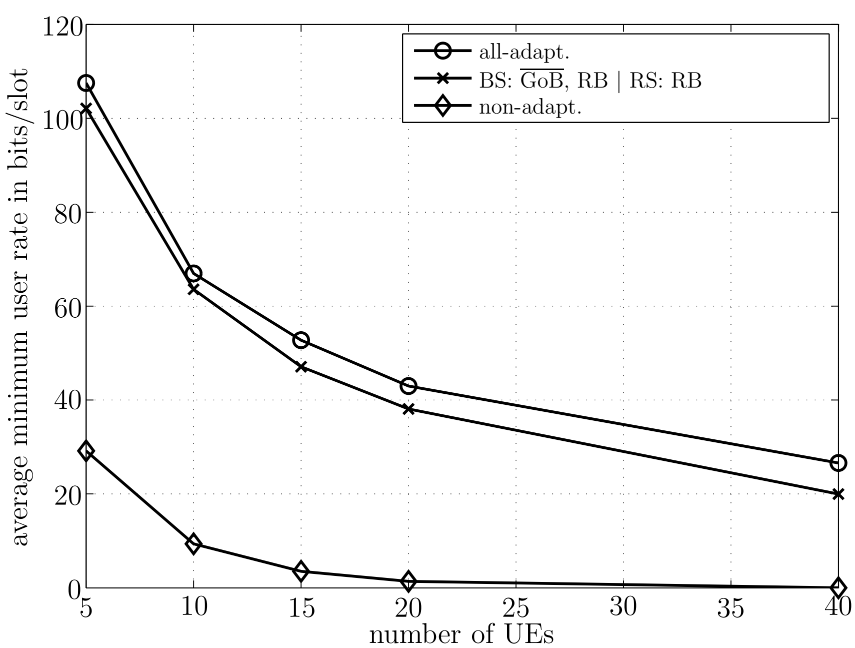

Figure 6.29. Comparison of the adaptive and non-adaptive algorithms for the design of grids of beams, $G_t = 3$ for all APs.

6.3.3.2 Maximizing the Sum Rate

In this section, the impact of the adaptive algorithm for the design of grids of beams is evaluated if the objective is the maximization of the sum rate. In Fig. 6.30, the average sum rate depending on the number of beams applied by the APs is illustrated. The results are given for the all-adaptive instance (all-adapt.), the instance using the non-adaptive algorithm for the design of grids of beams (BS: $\overline{\mathrm{GoB}}$, RB | RS: RB) and the non-adaptive instance (non-adapt.). The relative frequency of outages is close to zero except for the non-adaptive instance. Again, the results of the instance using the non-adaptive algorithm is not a smooth function of the number of beams since beams are chosen for grids of beams without taking into account co-channel interference. The adaptive algorithm for the design of grids of beams outperforms the non-adaptive one except for six beams. In the case of six beams, the same effect occurs as in Fig. 6.28. So many beams are applied that co-channel and inter-beam interference cannot be mitigated by the adaptive algorithm for the design of grids of beams better than mitigated by the non-adaptive algorithm.

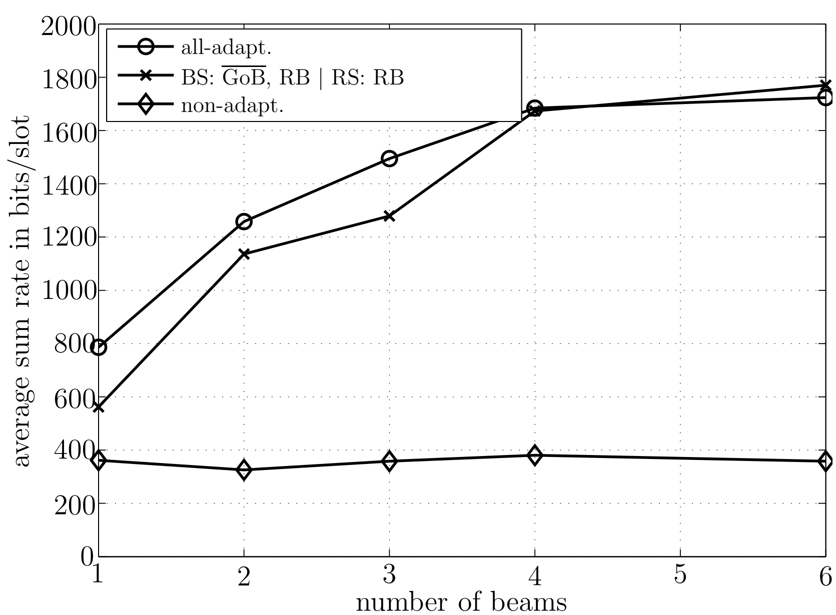

Figure 6.30. Comparison of the adaptive and non-adaptive algorithms for the design of grids of beams $N_{\mathrm{UE}} = 20$, $R_{\min} = 10$ bits/slot.

In Fig. 6.31, the average sum rate is shown depending on the number of UEs. The results of the all adaptive instance and the instance using the non-adaptive algorithm for the design of grids of beams show the same characteristic as already observed in Fig. 6.26. The all-adaptive instance and the instance using the non-adaptive

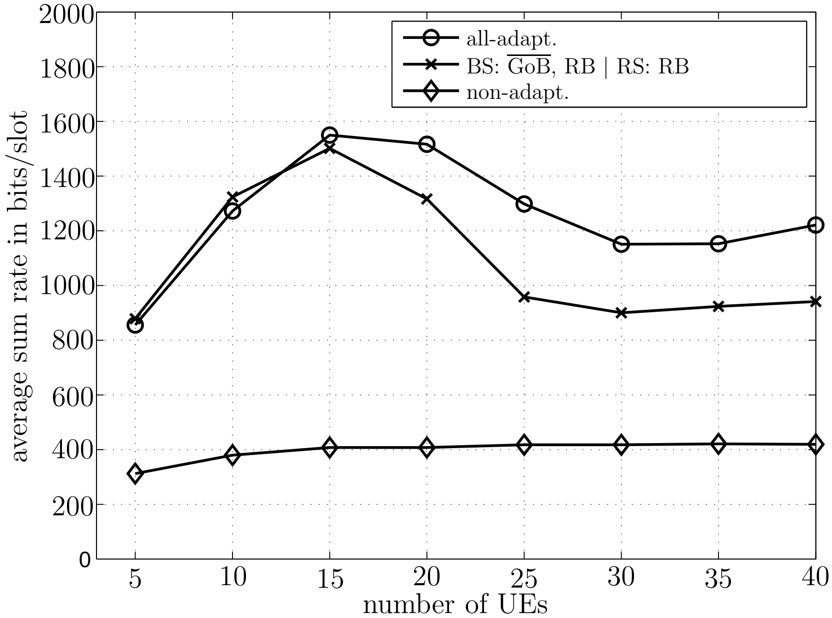

Figure 6.31. Comparison of the adaptive and non-adaptive algorithms for the design of grids of beams, $G_t = 3$ for all APs, $R_{\min} = 10$ bits/slot.

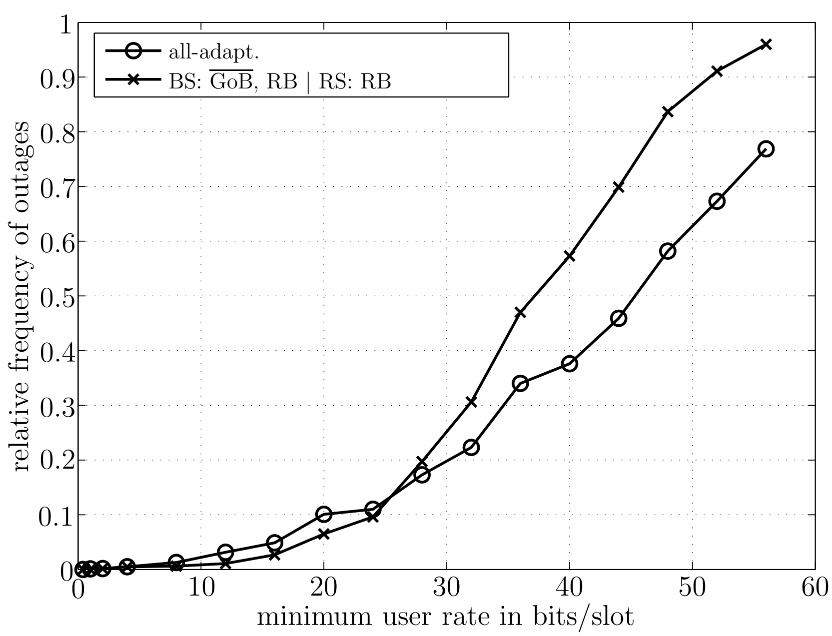

Figure 6.32. Comparison of the adaptive and non-adaptive algorithms for the design of grids of beams depending on the minimum user rate.

algorithm ensure an relative frequency of outages smaller than 5% if not more than 40 UEs are in the cell. For small numbers of UEs, the adaptive and non-adaptive algorithm for the design of grids of beams perform rather similar. If the number of UEs becomes large, the adaptive algorithm benefit from the demand values determined based on the PDF.

In Fig. 6.32, the relative frequency of outages is depicted as a function of the minimum user rate. If each UE each support with a small user rate, the all-adaptive instance and the instance using the non-adaptive algorithm behave rather similar. If the minimum user rate is large, the adaptive algorithm is preferable. This is in line with previous results, e.g., presented in Section. 6.2.3.2. If there are many UEs, more resources must be allocated among all UEs fairly and less resource blocks are available for the allocation to the small number of best UEs.

6.3.4 Allocation of Resource Blocks

6.3.4.1 Maximizing the Minimum User Rate

In this section, the impact of the algorithms allocating the resource blocks are analyzed if the distributed concept for reuse medium access is applied with the objective of the maximization of the minimum user rate. Three instances are considered: the all-adaptive instance (all-adapt.), the non-adaptive instance (non-adapt.) and the instance in which the allocation of resource blocks is solved by the RSs non-adaptively and the other subproblems are solved by the BS adaptively (BS: GoB, RB | RS: $\overline{\text{RB}}$). In Fig. 6.33 and Fig. 6.34, the average minimum user rate is depicted as a function of the number of beams applied by the APs and the number of UEs, respectively. The results reveal a huge gain if the RSs solve the allocation of resource blocks adaptively instead of non-adaptively. If the non-adaptive algorithm is applied, a UE is allocated a resource blocks defined by a beam that is not steered in the direction of the UE or affected by strong interference. This leads to a strong degradation of the corresponding SINR value and the number of bits carried by the resource block. It is concluded that the BS and the RSs shall always allocate the resource blocks adaptively.

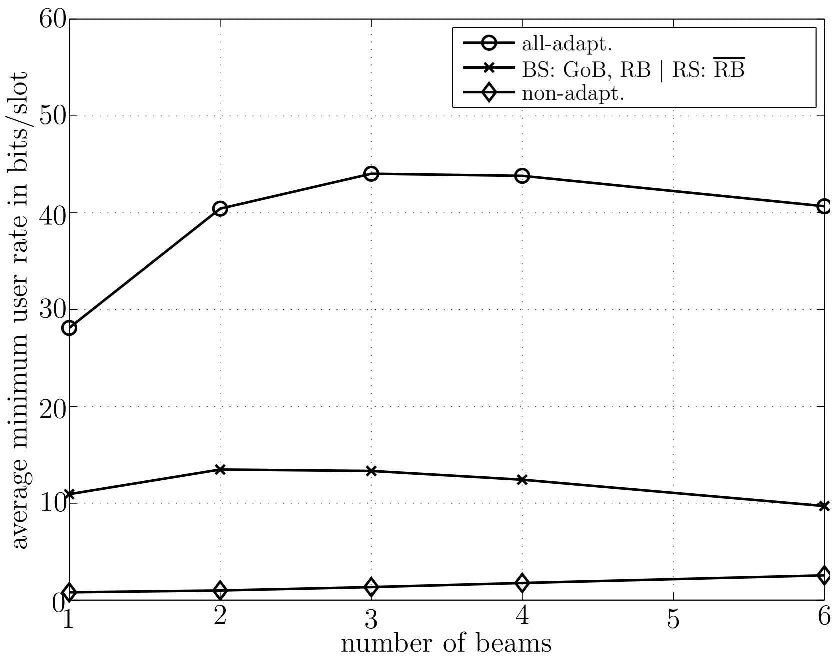

Figure 6.33. Comparison of various instances of adaptive and non-adaptive algorithms, $N_{\mathrm{UE}} = 20$.

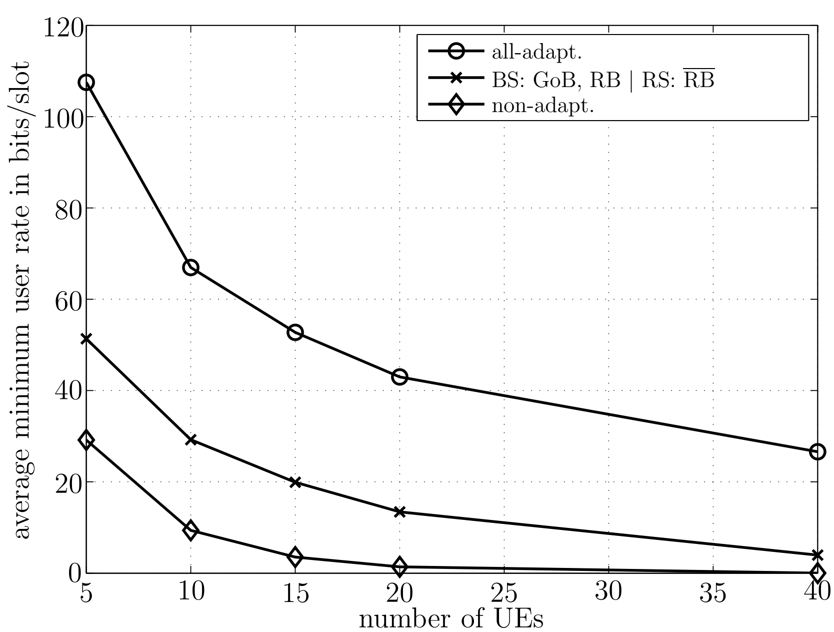

Figure 6.34. Comparison of various instances of adaptive and non-adaptive algorithms, $G_t = 3$ for all APs.

6.3.4.2 Maximizing the Sum Rate

In this section, the impact of the algorithms allocating the resource blocks are analyzed if the distributed concept for reuse medium access is applied with the objective of the maximization of the sum rate. The instances considered in this sections result from the same combinations of adaptive and non-adaptive algorithms as in Section 6.3.4, but the algorithms are changed with respect to the aim of the maximization of the sum rate. In Fig. 6.35, Fig. 6.36 and Fig. 6.37, the relative frequency of outages is depicted as a function of the number of beams applied by the APs, the number of UEs and the minimum user rate, respectively. The relative frequency of outages is considered since the minimum user rate cannot be provided rather often if the allocation of resource blocks is solved non-adaptively. This expectation is approved by all results. It is concluded that the BS and RSs shall always allocate the resource blocks adaptively.

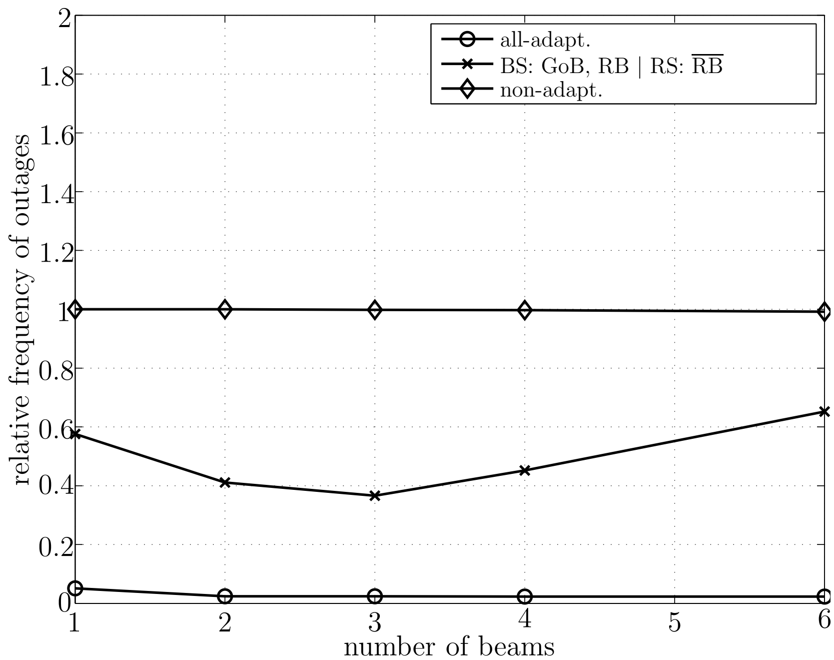

Figure 6.35. Comparison of various instance, $N_{\text{UE}} = 20$, $R_{\min} = 10$ bits/slot.

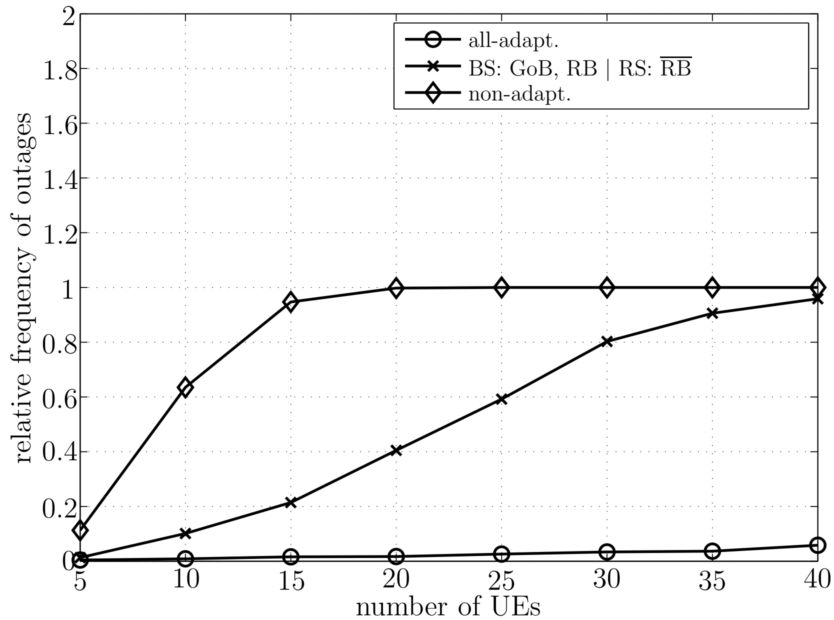

Figure 6.36. Comparison of various instances, $G_t = 3$ for all APs, $R_{\min} = 10$ bits/slot.

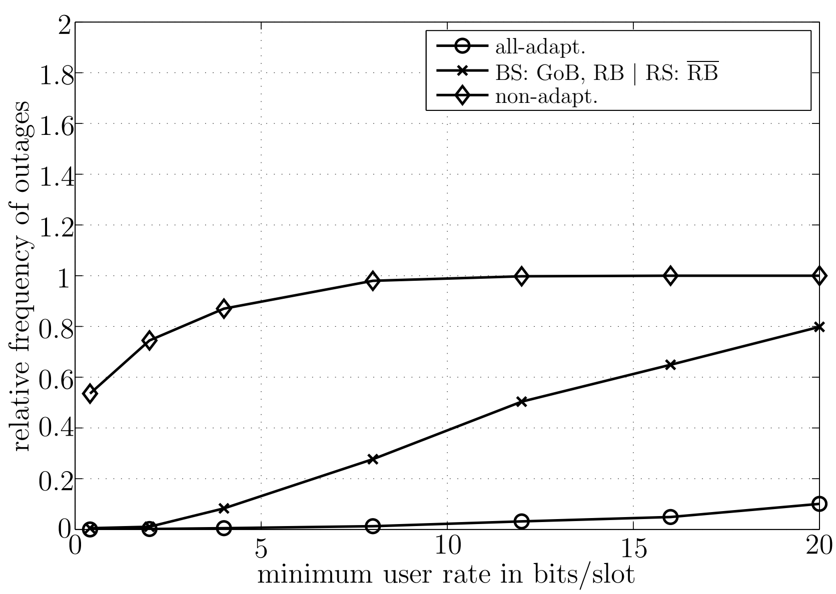

Figure 6.37. Comparison of various instances, $N_{\text{UE}} = 20$, $G_t = 3$ for all APs.

6.3.5 Summary for the Distributed Concept for Reuse Medium Access

In Section 6.3.2, the following results are derived:

- The distributed concept for reuse medium access enables adaptive algorithms for the design of grids of beams and the allocation of resource blocks. The subproblems are solved based on a signalling overhead from RS to BS which is much lower than required for a solution at a central unit of the cell.
- If the objective is defined by (P1), the following gains compared to non-adaptive algorithms are achieved:
  - The adaptive algorithm for the design of grids of beams improves the average minimum user rate slightly if the number of beams applied by the APs is rather small.
  - The adaptive algorithms applied by the BS and RSs for the allocation of resource blocks achieve is a must. A non-adaptive algorithm leads to a low average minimum user rate.
- If the objective is defined by (P2), the following gains compared to non-adaptive algorithms are achieved:
  - The adaptive algorithm for the design of grids of beams improves the average sum rate for a large number of UEs in the cell and improves the relative frequency of outages for a large minimum user rate.
  - Resource blocks must be allocated adaptively, as already stated for the objective (P1) since a huge gain in terms of average sum rate and relative frequency of outages is achieved.

Chapter 7
Conclusions

This thesis deals with the allocation of resources, namely beams, resource blocks, power and slots in the cell of an OFDMA-based relay network. Two objectives are defined for the allocation of resources. The first one aims at the maximization of the minimum user rate and corresponds to a totally fair allocation in terms of equal user rates among all UEs in the cell. The second one aims at the maximization of the sum rate. Since the pure maximization of the sum rate does not consider any fairness among UEs, it is claimed that a minimum user rate must be provided to each UE.

Chapter 2 provides the system model required to describe resource allocation problems in the cell of a relay network. The system model is applicable to two types of scenarios differing in the medium access required to organize the transmissions of multiple APs. The first one models scenarios in which the user rates are limited by noise since an orthogonal medium access is considered. The second one models scenarios in which the user rates are limited by co-channel interference since reuse medium access is considered.

Based on this system model, two resource allocation problems called (P1) and (P2) are defined in Chapter 3. The definitions are related to the two objectives of the maximization of the minimum user rate and the maximization of the sum rate. It is motivated that a optimum solution of the problems is impractical due to limitations concerning computational complexity and signalling overhead between RSs and BS. The distributed concept for orthogonal medium access and the distributed concept for reuse medium access are introduced. The former concept is designed for the orthogonal medium access and is based on channel gain values. The latter one is designed for the reuse medium access and is based on SINR values. Each concept is applicable to both objectives. A concept decomposes a considered resource allocation problem in smaller subproblems such that a lower computational complexity is required to solve the subproblems. The subproblems are partly solved by the BS and partly by the RSs in order to distribute the required computational complexity. Additionally, the signalling overhead sent from the RSs to the BS is kept low since initial data rate values are defined.

The subproblems related to the distributed concept for orthogonal medium access are defined in Chapter 4. The subproblems are the design of grids of beams, the allocation of resource blocks, the allocation of power and bits and the allocation of slots. Novel algorithms enabling an adaptive allocation at a low computational complexity are presented for these subproblems. Except for the allocation of slots, non-adaptive algorithms are presented, too. Although, the adaptive algorithms are expected to perform better in terms of the objective defined in (P1) and (P2), the adaptive algorithms are only applied, if the computational complexity can be fulfilled. If this is not the case, the concept allows to replace an adaptive algorithm by the corresponding non-adaptive algorithm. The allocation of slots is excluded since a non-adaptive algorithm affects the allocation of resource blocks, power and bits. This is treated in the distributed concept for reuse medium access.

The subproblems related to the distributed concept for reuse medium access are defined in Chapter 5. The subproblems are the allocation of slots, the design of grids of beams and the allocation of resource blocks. Novel algorithms enabling an adaptive allocation at a low complexity are proposed except for the allocation of slots. A fixed number of slots is always allocated to the APs in order to gain the SINR values required to solve the allocation of resource blocks. For all subproblems, non-adaptive algorithms are presented, too.

The concepts including the introduced algorithms are evaluated in Chapter 6. In an exemplary scenario, it is shown that the concepts are based on a signalling overhead from a RS to the BS which is much lower than required for an optimum solution of (P1) and (P2). Additionally, the adaptive algorithms show strong performance gains compared to the non-adaptive ones. The computational complexity required to apply all adaptive algorithms can be offered by today's processors. Hence, the concepts allow an applicable and efficient allocation of resources in an OFDMA-based relay network.

List of Acronyms

AP	Access Point
BS	Base Station
CQI	Channel Quality Information
CSI	Channel State Information
FDD	Frequency Division Duplex
FFT	Fast Fourier Transform
FLOPS	Floating Point Operations Per Second
GSM	Global System for Mobile communications
HIPERLAN	HIgh PErformance Radio Local Area Network
IEEE	Institute of Electrical and Electronics Engineers
IFFT	Inverse Fast Fourier Transform
LOS	Line-Of-Sight
LTE	Long Term Evolution
NLOS	Non-Line-Of-Sight
NP-hard	Nondeterministic Polynomial time hard
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PDF	Probability Density Function
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RS	Relay Station
SDMA	Spatial Division Multiple Access
SINR	Signal-to-Interference plus Noise Ratio

SNR	Signal-to-Noise Ratio
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
UCA	Uniform Circular Array
UE	User Equipment
UMTS	Universal Mobile Telecommunications System

List of Symbols

A_t	Coverage area of AP t
$A^{(t,b)}$	Coverage area of AP t and beamforming vector b
$\acute{\mathbf{A}}$	Interference matrix for time-frequency unit
$\mathbf{A}$	Interference matrix for all time-frequency units
$\mathbf{A}_f$	Matrix of the product of channelgain values and SINR thresholds
b	Index of available beamforming vector
B_t	Number of beamforming vectors available at AP t
$BEP_{t,r}$	Bit error probability on link (t, r)
$c_{t,k}$	Modulation symbol on resource block k allocated by AP t
$D_{t,b}$	Demand value for beamforming vector b of AP t
$\mathbf{d}_{r,k}$	Receive filter vector of receiving station r for resource block k
$\mathrm{E}\{\cdot\}$	Expectation operator
f	Index of time-frequency unit
F	Number of time-frequency units in a slot
$\mathbf{f}_s$	Vector of number of bits for slot s
G_t	Number of beams applied to a time-frequency unit
$G_{(t,b)}$	Antenna gain of beamforming vector b and AP t in interference modell
$h^{(i,j)}_{t,r,f}$	Channel coefficient of the antenna pair (i, j) of the link (t, r) on time-frequency unit f
$\mathbf{H}_{t,r,f}$	Matrix of channel coefficients of the link (t, r) on time-frequency unit f
i	Index of transmit antenna
$I_{\mathrm{CCI},r,k}$	Co-channel interference power received by receiving station r on resource block k
$I_{\mathrm{IBI},r,k}$	Inter-beam interference power received by receiving station r on resource block k
$I_{(t,b)}$	Interference generated by beamforming vector b of AP t
$I^{(t,b)}_{(t',b')}$	Average interference generated by beamforming vector b' of AP t' and in the sector of beamforming vector b and AP t
j	Index of receive antenna
k	Index of resource block
K_t	Number of resource blocks available for the transmitting AP t in a slot
$\mathbf{m}_{t,b}$	b-th beamforming vector of AP t

$\mathbf{m}_{t,k,n}$	Beamforming vector of AP t applied for resource block k in subframe n
n	Subframe index
$N_{\text{DC-OMA}}$	Number of bits per frame fed back from a RS to the BS for a distributed concept
N_{Max}	Number of bits per frame fed back from a RS to the BS for a central solution
N_r	Number of antennas of receiving station r
N_{RS}	Number of RSs in the cell
N_{SC}	Number of subcarriers in a time-frequency unit
N_{SF}	Number of subframes in a frame
N_t	Number of antennas of AP t
N_{UE}	Number of UEs in cell
N_{Val}	Number of values fed back from RS to BS
p	Sum of allocated power
$p_{t,k}$	Power allocated by AP t to the k-th resource block
$P_{r,k}$	Signal power received by receiving station r for resource block k
P_t	Maximum power of AP t
$P^{\text{TX}}_{(t,b)}$	Transmit power component of beamforming vector b and AP t in interference modell
PL_t	Path loss component of AP t in interference modell
Pr_t	Probability that a UE is in the coverage area of AP t
$\mathbf{p}_f$	Vector of power values for time-frequency unit f
r	Index of receiving station
$R_{0,r',r}$	Data rate addressed to UE r and transmitted from BS to RS r'
$R_{\min,t,r}$	Minimum data rate for link (t,r)
$R_{t,r}$	Data rate on link (t,r)
$\hat{R}_{t,r}$	Initial data rate value of link (t,r)
$R_{\sum}$	Sum rate
s	Slot index
$\mathbf{s}_s$	Vector of state variables for slot s
S	Number of slots in a frame
S_n	Number of slots in subframe n
t	Index of transmitting station
$u_{t,r,k,\epsilon}$	Assignment variable indicating the allocation of a resource block k to link (t,r) with ϵS_n bits in subframe n

$v_{t,b,k,n}$	Assignment variable indicating the use of beam b for resource block k by AP t in subframe n
$\mathbf{v}$	Vector of assignment variable $v_{t,b,k,n}$
$w_{r,k}$	weight of receiver r and resource block k
x	Cartesian coordinate
$\mathbf{x}_s$	Vector of decision variables for slot s
$\mathbf{x}_{t,k}$	Transmit signal vector of AP t for resource block k
y	Cartesian coordinate
$\hat{y}_{r,k}$	Received signal of receiver r for resource block k
$\mathbf{y}_{r,k}$	Received signal of receiver r for resource block k before receive filter
Z_0	Average noise power
$Z_{r,k}$	Noise power on the k-th resource block of receiving station r
$\mathbf{z}_{r,k}$	Noise vector received on the f-th time-frequency unit by receiver r
$\mathbf{Z}_f$	Vector of noise values of time-frequency unit f
$\alpha^2_{t,r,k,n}$	Channelgain of resource block k on link (t, r) in subframe n
$\gamma_{r,k}$	SINR value of receiving station r on resource block k
$\gamma_{\epsilon,r}$	SINR value required to transmit ϵ bits to receiving station r
Δ	Difference of data rate values of the BS-to-RS and RS-to-UE link
$\Delta\mathbf{p}_k$	Vector of incremental power values for resource block k
$\Delta\epsilon$	Increment of number of bits represented by a modulation symbol
ϵ	Number of bits represented by a modulation symbol
$\epsilon_{r,k}$	Number of bits allocated to a slot of resource block k addressed to receiving station r
ρ	PDF of user distribution in the cell
$\mathcal{E}$	Set of number of bits represented by a modulation symbol
$\mathcal{F}_t$	Set of time-frequency units of AP t
$\mathcal{G}_{t,f}$	Set of beams allocated to time-frequency unit f of AP t
$\mathcal{K}$	Set of resource blocks
$\mathcal{K}_r$	Set of resource blocks allocated to receiving station r
$\mathcal{R}_{\mathrm{RS}}$	Set of relay stations
$\mathcal{R}_{\mathrm{RS},b}$	Set of relay stations in direction of beam b
$\mathcal{R}_t$	Set of stations receiving from t
$\mathcal{T}_n$	Set of reuse group given by the transmitters of a cell using subframe n
$\mathbb{C}$	Set of complex numbers
$\mathbb{R}$	Set of real numbers

$\lvert \cdot \rvert$	Cardinality of a set
$\lvert\lvert \cdot \rvert\rvert$	Absolute value of a scalar or 2-norm of a vector
$(\cdot)^{\mathrm{H}}$	Conjugate transpose of a vector or matrix
$(\cdot)^{*}$	Conjugate complex
$(\cdot)^{\mathrm{T}}$	Transpose of a vector or matrix

Bibliography

[3GP00] 3GPP, "Digital cellular telecommunications system (Phase 2+); radio transmission and reception," Technical Specification Group GERAN, Tech. Rep. TR 45.005 V4.0.0, Oct. 2000.

[3GP06] ——, "Evolved universal terrestrial radio access (E-UTRA) and evolved universial terrestrial radio access network (E-UTRAN); radio interface protocol aspects," Technical Specification Group Radio Access Network, Tech. Rep. TR 25.813 V7.1.0, Sep. 2006.

[3GP09] ——, "Technical specifications and technical reports for a UTRAN-based 3GPP system," Technical Specification Group Services and System Aspects, Tech. Rep. TS 21.101 V8.0.0, Mar. 2009.

[BGMW07] M. Bohge, J. Gross, M. Meyer, and A. Wolisz, "Dynamic resource allocation in OFDM systems: An overview of cross-layer optimization principles and techniques," *IEEE Network Magazine*, vol. 21, pp. 53–59, Jan. 2007.

[BHIT05] C. F. Ball, E. Humburg, K. Ivanov, and F. Treml, "Performance analysis of IEEE 802.16 based cellular MAN with OFDM-256 in mobile scenarios," *in Proc. of 63rd IEEE Vehicular Technology Conference*, Jun. 2005.

[Bla84] R. E. Blahut, *Theory and Practice of Error Control Codes.* Addison-Wesley, 1984.

[Bos98] M. Bossert, *Kanalcodierung*, 1st ed. B.G. Teubner Stuttgart, 1998.

[BV04] S. Boyd and L. Vandenberghe, *Convex optimization*, 1st ed. Cambridge University Press, 2004.

[BYFP04] H. Bolukbasi, H. Yanikomeroglu, D. D. Falconer, and S. Periyalwar, "On the capacity of cellular fixed relay networks," *in Proc. of Canadian Conference on Electrical and Computer Engineering*, May 2004.

[CCB95] P. Chow, J. Cioffi, and J. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Transactions on Wireless Communications*, vol. 43, pp. 773–775, Feb. 1995.

[CFRZ07] E. Costa, A. Frediani, S. Redana, and Y. Zhang, "Dynamic spatial resource sharing in relay enhanced cells," *in Proc. of European Wireless*, Apr. 2007.

[CFRZ08] ——, "Dynamic spatial resource sharing in relay enhanced cells," *in Proc. of European Wireless*, Apr. 2008.

[CG79] T. M. Cover and A. A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Transaction on Information Theory*, vol. 25, pp. 572–584, Sep. 1979.

[CH04] J. Cho and Z. J. Haas, “On the throughput enhancement of the downstream channel in cellular radio networks through multihop relaying,” *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 1206–1219, Sep. 2004.

[CKL+08] L. Carrington, D. Komatitsch, M. Laurenzano, M. Tikir, D. Michea, N. L. Goff, A. Snavely, and J. Tromp, “High-frequency simulations of global seismic wave propagation using SPECFEM3D GLOBE on 62K processors,” *In Proc. of International Conference for High Performance Computing, Networking, Storage and Analysis*, Nov. 2008.

[CZT07] E. Costa, Y. Zhang, and A. Totaro, “Dynamic resource allocation in relay enhanced cells based on winner system,” *in Proc. of 16th IST Mobile and Wireless Communications Summit*, Jul. 2007.

[Doh04] M. Dohler, “Resource allocation for FDMA-based regenerative multihop links,” *IEEE Transactions on Wireless Communications*, vol. 3, Nov. 2004.

[DSA02] M. Dohler, F. Said, and H. Aghvami, “Concept of virtual antenna arrays,” *in Proc. of IEEE Globecom 2002*, Dec. 2002.

[DXWS07] K. Doppler, H. Xiaoben, C. Wijting, and A. Sorri, “Adaptive soft reuse for relay enhanced cells,” *in Proc. of IEEE 65th Vehicular Technology Conference 2007*, Apr. 2007.

[DZC99] L. Devroy and C. Zamora-Cura, “On the complexity of branch-and bound search for random trees,” *Random Structures and Algorithms*, vol. 14, pp. 309–327, Jul. 1999.

[ECV03] M. Ergen, S. Coleri, and P. Varaiya, “Qos aware adaptive resource allocation techniques for fair scheduling in OFDMA based broadband wireless access systems,” *IEEE Transactions on Broadcast*, vol. 49, pp. 362–370, May 2003.

[ETS97] ETSI, “Broadband radio access networks (BRAN), HIgh PErformance Radio Local Area Network (HIPERLAN),” Type 2; Requirements and architectures for wireless broadband access, Tech. Rep. TR 101 031 V7.1.0, Apr. 1997.

[ETS03] ——, “Fixed radio systems; point-to-multipoint equipment; time division multiple access (TDMA); point-to-multipoint digital radio systems in the range of 3 GHz to 11 GHz,” Tech. Rep. EN 301 021 V1.4.1, Jul. 2003.

[EVW01] N. Esseling, H. Vandra, and B. Walke, “A forwarding concept for HiperLAN/2,” *Computer Networks*, vol. 37, pp. 25–32, Sep. 2001.

[FBB+08] I. Frigyes, P. Bakki, J. Bito, T. F. Maciel, and A. Klein, *Advances in Mobile and Wireless Communications - Views of the 16th IST Mobile and Wireless Communication Summit*, 1st ed. Springer, 2008.

[For73] D. Forney, “The viterbi algorithm,” *Proceedings of the IEEE*, vol. 61, pp. 268–278, Mar. 1973.

[FT07] G. Ferrari and O. Tonguz, “Impact of mobility on the BER performance of ad hoc wireless networks,” *IEEE Transactions on Vehicular Technology*, vol. 56, pp. 271–286, Oct. 2007.

[FY05] A. Florea and H. Yamikomeroglu, “On the optimum number of hops in infrastructure-based fixed relay networks,” *in Proc. of IEEE Globecom 2005*, Dec. 2005.

[GRB06a] R. Grünheid, H. Rohling, and K. Brüninghaus, “Self-organised beamforming and opportunistic scheduling in an OFDM-based cellular network,” *in Proc. of 63rd IEEE Vehicular Technology Conference*, Apr. 2006.

[GRB06b] ——, “Self-organised beamforming and opportunistic scheduling in an OFDM-based cellular network,” *in Proc. of 63rd IEEE Vehicular Technology Conference*, Apr. 2006.

[HH87] D. Hughes-Hartogs, “Ensemble modem structure for imperfect transmission media,” *US Patent 4679227*, Jul. 1987.

[HK06] C. Hoymann and K. Klagges, “MAC frame concepts to support multihop communication in IEEE 802.16 networks,” *in Proc. of 16th Wireless World Research Forum*, Apr. 2006.

[HN04] T. E. Hunter and A. Nosratinia, “Diversity through coded cooperation,” *IEEE Transactions on Wireless Communications*, vol. 5, pp. 283–289, Feb. 2004.

[Hoy04] C. Hoymann, “Analysis and performance evaluation of the OFDM-based metropolitan area network ieee 802.16,” *Computer Networks*, vol. 49, pp. 341–363, Jun. 2004.

[HR08] A. T. H. Rohling, R. Grünheid, “Beamforming and scheduling in a cellular OFDM system,” *in Proc. of International ITG/IEEE Workshop on Smart Antennas*, Mar. 2008.

[HRW+07] L. Huang, M. Rong, L. Wang, Y. Xue, and E. Schulz, “Resource scheduling for OFDMA/TDD based relay enhanced cellular networks,” *in Proc. of Wireless Communications and Networking Conference*, Mar. 2007.

[Hun07] R. Hunger, “Floating point operations in matrix-vector calculus,” *Munich University of Technology, Tech. Rep. v1.3*, Sep. 2007.

[IEE04] IEEE Standard 802.16, “Part 16: Air interface for fixed broadband wireless access systems,” Tech. Rep., Oct. 2004.

[IEE07] IEEE Standard 802.11, “Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications,” Tech. Rep., Jun. 2007.

[IEE08] IEEE Draft Standard 802.16j/D5, "Part 16: Air interface for fixed and mobile broadband wireless access systems - multihop relay specification," Tech. Rep., May 2008.

[IST] IST WINNER Website. [Online]. Available: http://www.ist-winner.org

[IST07a] IST-4-027756 WINNER II, "Final assessment of relaying concepts for all CGs scenarios under considerations of related WINNER II L1 and L2 protocol functions," Tech. Rep. Technical Report D3.5.3 v1.0, Sep. 2007. [Online]. Available: www.ist-winner.org

[IST07b] ——, "WINNER II channel models," Tech. Rep. Technical Report D1.1.2 v1.0, Sep. 2007. [Online]. Available: www.ist-winner.org

[JH06] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 3524–3536, Dec. 2006.

[JHHN04] M. Janani, A. Hedayat, T. E. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: Space-time transmission and iterative decoding," *IEEE Transactions on Signal Processing*, vol. 52, pp. 362 – 371, Feb. 2004.

[JUN05] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Transactions on Signal Processing*, vol. 8, pp. 2700–2712, Aug. 2005.

[Ken88] D. Kennedy, "Some branch and bound techniques for nonlinear optimization," *Mathematical Programming*, vol. 42, pp. 147–157, Apr. 1988.

[KF08] C. Koutsimanis and G. Fodor, "A dynamic resource allocation scheme for guaranteed bit rate services in OFDMA networks," *in Proc. of IEEE International Conference on Communications 2008*, May 2008.

[KGG05] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, pp. 3037–3063, Sep. 2005.

[KPL06] I. Kim, I.-S. Park, and Y. H. Lee, "Use of linear programming for dynamic subcarrier and bit allocation in multiuser OFDM," *IEEE Transaction on Vehicular Technology*, vol. 55, pp. 1195–1207, Jul. 2006.

[LCS01] X. Liu, E. K. P. Chong, and N. B. Shroff, "Opportunistic transmission scheduling with resource-sharing constraints in wireless network," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 2053–2064, Oct. 2001.

[LHYT06] Y. Liu, R. Hoshyar, X. Yang, and R. Tafazolli, "Integrated radio resource allocation for multihop cellular networks with fixed relay stations," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 2137–2146, Nov. 2006.

[Lin03] J. C. Lin, “Health aspects of wireless communication: personal wireless communication radiation and the eye lens,” *in Proc. of ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 7, pp. 4–7, Jul. 2003.

[LL00] B. K. Lau and Y. H. Leung, “A dolph-chebyshev approach to the synthesis of array patterns of uniform linear arrays,” *in Proc. of IEEE International Symposium on Circuits and Systems*, May 2000.

[LL05] H. Liu and G. Li, *OFDM-Based Broadband Wireless Networks - Design and Optimization*, 1st ed. Wiley, 2005.

[LL06] K.-D. Lee and V. Leung, “Fair allocation of subcarrier and power in an OFDMA wireless mesh network,” *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 2051–2060, Nov. 2006.

[LLW+02] H. Li, M. Lott, M. Weckerle, W. Zirwas, and E. Schulz, “Multihop communications in future mobile radio networks,” *in Proc. of 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep. 2002.

[LTW04] J. Laneman, T. N. C. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Transactions on Information Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[LW03] J. N. Laneman and G. W. Wornell, “Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks,” *IEEE Transactions on Information Theory*, vol. 49, pp. 2415–2425, Oct. 2003.

[LZ06] K. B. Letaif and Y. Z. Zhang, “Dynamic multiuser resource allocation and adaptation for wireless systems,” *IEEE Wireless Communications*, vol. 13, pp. 38–47, Aug. 2006.

[Mac09] T. F. Maciel, “Suboptimal resource allocation for multi-user MIMO-OFDMA systems,” Ph.D. dissertation, Technische Universität Darmstadt, 2009.

[MKR09] C. Müller, A. Klein, and B. Raaf, “A coordinated resource allocation algorithm for infrastructure-based relay networks,” *EURASIP Journal on Advances in Signal Processing*, Jan. 2009.

[MKW06] C. Müller, A. Klein, and F. Wegner, “Coverage extension of WiMax using multihop in a low user density environment frequenz,” *in Proc. of International OFDM-Workshop*, Aug. 2006.

[MKW07] ——, “Coverage extension of WiMax using multihop in a low user density environment frequenz,” *Frequenz, Journal of RF-Engineering and Telecommunications*, vol. 61, pp. 2–6, Jan. 2007.

[MKWK07a] C. Müller, A. Klein, F. Wegner, and M. Kuipers, “Costs and performance of non-cooperative relay networks,” *in Proc. of European Wireless*, Apr. 2007.

[MKWK07b] ——, “Dynamic subcarrier, bit an power allocation in OFDMA-based relay networks,” *in Proc. of International OFDM-Workshop*, Aug. 2007.

[MSJF07] P. Moberg, P. Skillermark, N. Johansson, and A. Furuskär, “Performance and cost evaluation of fixed relay nodes in future wide area cellular networks,” *in Proc. of IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep. 2007.

[NP00] R. Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*, 1st ed. Artech House Publishers, 2000.

[NW99] G. Nehmhauser and L. Wolsey, *Integer and Combinatorial Optimization.* Wiley, 1999.

[Pap84] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 2nd ed. McGraw-Hill Book Company, 1984.

[Par00] J. D. Parsons, *The Mobile Radio Propagation Channel*, 2nd ed. Wiley, 2000.

[PF05] D. P. Palomar and J. R. Fonollosa, “Practical algorithms for a family of waterfilling solutions,” *IEEE Transactions on Signal Processing*, vol. 53, pp. 687–695, Feb. 2005.

[PIPNL04] A. Pascual-Iserte, A. Prez-Neira, and M. A. Lagunas, “An approach to optimum joint beamforming design in a MIMO-OFDM multiuser system,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2004, pp. 210–221, 2004.

[PNG03] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*, 1st ed. Cambridge University Press, 2003.

[Pro95] J. G. Proakis, *Digital Communications*, 3rd ed., ser. Electrical Engineering Series. McGraw-Hill, 1995.

[PWS+04] R. Pabst, B. Walke, D. Schultz, P. Herold, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. Falconer, and G. Fettweis, “Relay-based deployment concepts for wireless and mobile broadband radio,” *IEEE Communications Magazine*, vol. 42, pp. 80–89, Sep. 2004.

[Qiu04] W. Qiu, “Transmit power reduction in MIMO multi-user mobile radio downlinks by the rationale receiver orientation,” Ph.D. dissertation, Universität Kaiserslautern, 2004.

[RC00] W. Rhee and J. Cioffi, “Increase capacity of multiuser OFDM system using dynamic subchannel allocation,” *in Proc. of 51st IEEE Vehicular Technology Conference*, May 2000.

[RGT08] H. Rohling, R. Grünheid, and A. Tassoudji, “Beamforming and scheduling in a cellular OFDM system,” *in Proc. of International ITG/IEEE Workshop on Smart Antennas*, Feb. 2008.

[RW07] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 379–389, Feb. 2007.

[SAR09] S. Sadr, A. Anpalagan, and K. Raahemifar, "Radio resource allocation algorithms for the downlink of multiuser ofdm communication systems," *IEEE Communications Surveys and Tutorials*, vol. 11, pp. 92–106, 2009.

[SAY06] H. Shi, T. Asai, and H. Yoshino, "A relay node division duplex relaying approach for MIMO relay networks," *in Proc. of IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep. 2006.

[SBS99] L. Songwu, V. Bharghavan, and R. Srikant, "Fair scheduling in wireless packet networks," *IEEE/ACM Transactions on Networking*, vol. 7, pp. 473–489, Aug. 1999.

[SE04] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Transactions on Communications*, vol. 52, pp. 1470 – 1476, Sep. 2004.

[SEW07] R. Schoenen, J. Eichinger, and B. Walke, "On the OFDMA FDD mode in 3G-LTE," *in Proc. of International OFDM-Workshop*, Aug. 2007.

[SS04] Q. H. Spencer and A. L. Swindlehurst, "Channel allocation in multi-user MIMO wireless communications systems," *in Proc. of IEEE International Conference on Communications*, Jun. 2004.

[SSH04] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, pp. 461–471, Feb. 2004.

[ST81] W. L. Stutzmann and G. A. Thiele, *Antenna Theory and Design*, 1st ed. John Wiley & Sons, 1981.

[SWPI03] D. C. Schultz, B. Walke, R. Pabst, and T. Irnich, "Fixed and planned relay based radio network deployment concepts," *in Proc. of 10th Wireless World Research Forum*, Oct. 2003.

[SZQY05] J. Shi, Z. Zhang, P. Qiu, and G. Yu, "Subcarrier and power allocation for OFDMA-based regenerative multi-hop links," *in Proc. of International Conference on Wireless Communications, Networking and Mobile Computing*, Sep. 2005.

[Tan96] A. S. Tanenbaum, *Computer Networks*, 3rd ed. Prentice Hall, 1996.

[Tim05] B. Timus, "Cost analysis issues in a wireless multihop architecture with fixed relays," *in Proc. of 61st IEEE Vehicular Technology Conference*, May 2005.

[UMK06] T. Unger, C. Müller, and A. Klein, "Resource allocation schemes for infrastructure based relay," *in Proc. of International OFDM-Workshop*, Aug. 2006.

[VALP06] M. Vemula, D. Avidor, J. Ling, and C. Papadias, "Inter-cell coordination, opportunistic beamforming and scheduling," *in Proc. of IEEE International Conference on Communications*, Jun. 2006.

[vdM71] E. C. van der Meulen, "Three-terminal communication channels," *Advances in Applied Probability*, vol. 3, pp. 120–154, 1971.

[VTL02] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Transaction on Information Theory*, vol. 48, pp. 1277–1294, Jun. 2002.

[WCLM99] C. Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 1747–1758, Oct. 1999.

[WQDT01] H. Wu, C. Qiao, S. De, and O. Tonguz, "Integrated cellular and ad hoc relaying systems: iCAR," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 2105–2115, Oct. 2001.

[YSL06] S. Yiu, R. Schober, and L. Lampe, "Distributed space-time block coding," *IEEE Transactions on Wireless Communications*, vol. 54, pp. 1195–1206, Jul. 2006.

[ZHF04] E. Zimmermann, P. Herhold, and G. Fettweis, "A novel protocol for cooperative diversity in wireless networks," *in Proc. of 5th European Wireless Conference*, Feb. 2004.

[ZHF05] ——, "On the performance of cooperative relaying protocols in wireless networks," *European Transactions on Telecommunications*, vol. 16, pp. 5–16, Jan. 2005.

[ZL04] Y. J. Zhang and K. Letaief, "Multiuser adaptive subcarrier-and-bit allocation with adaptive cell selection for OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 1566–1575, Sep. 2004.

Lebenslauf

Name:	Ulrich Christian Müller
Anschrift:	Fichtenstraße 38, 65527 Niedernhausen
Geburtsdatum:	12.03.1979
Geburtsort:	Andernach

Schulausbildung

1985-1989	Grundschule in Mendig
1989-1998	Kurfürst-Salentin-Gymnasium in Andernach

Studium

1999-2004	Studium der Elektro- und Informationstechnik an der Technischen Universität Darmstadt, Studienabschluß: Diplom-Ingenieur

Berufstätigkeit

2004-2009	wissenschaftlicher Mitarbeiter am Fachgebiet Kommunikationstechnik, Institut für Nachrichtentechnik, Technische Universität Darmstadt